DEGREE PROJECT IN COMMUNICATION SYSTEMS, FIRST LEVEL
STOCKHOLM, SWEDEN 2015

Internet of Things

Exploring and Securing a Future Concept

CRISTIAN BUDE and ANDREAS KERVEFORS BERGSTRAND

KTH ROYAL INSTITUTE OF TECHNOLOGY
INFORMATION AND COMMUNICATION TECHNOLOGY

Internet of Things

Exploring and Securing a Future Concept

Cristian Bude and Andreas Kervefors Bergstrand

2015-06-15

Bachelor's Thesis

Examiner and Academic adviser
Gerald Q. Maguire Jr.

Industrial adviser
Emma Andersdotter and Johan Thulin

KTH Royal Institute of Technology
School of Information and Communication Technology (ICT)
Department of Communication Systems
SE-100 44 Stockholm, Sweden

Abstract

Internet of Things (IoT) is a concept that encompasses various objects and methods of communication to exchange information. Today IoT is more a descriptive term of a vision that everything should be connected to the internet. IoT will be fundamental in the future because the concept opens up opportunities for new services and new innovations. All objects will be connected and able to communicate with each other, while they operate in unprotected environments. This later aspect leads to major security challenges.

Today, IoT is in great need of standardization and clear architectures that describe how this technology should be implemented and how IoT devices interact with each other in a secure manner. The security challenges are rooted in the technology and how information is acquired and manipulated by this technology. This thesis provides an introduction to what the IoT is and how it can be used as well as some of the threats that IoT may face in regards to information security. In addition, the thesis provides the reader with some suggestions about how to potentially solve the fundamental need for authentication and secure communications. The solutions presented are based on both contemporary solutions and technologies that are under development for the future. Contemporary solutions are based on security protocols such as IPSec and DTLS. These protocols are being used in an environment that extends across the Internet and into a 6LoWPAN network. The proposed authentication solution has been developed based on a public key infrastructure and trust models for certificate management.

As future work, the thesis presents several research areas where this thesis can be used as a basis. These specialization areas include further analysis of vulnerabilities and an implementation of the proposed solutions.

Keywords:

Internet of Things, IoT, information security, identification, authentication, secure communication

Sammanfattning

Internet of Things (IoT) är ett koncept som omfattar olika objekt och kommunikationsmetoder för utbyte av information. Idag är IoT mer en beskrivande term av den framtidsvision som finns att allting ska vara uppkopplat på internet. IoT kommer vara fundamentalt i framtiden eftersom konceptet öppnar upp möjligheter för nya tjänster samt nya innovationer. Då alla objekt ska vara uppkopplade och kunna kommunicera med varandra samtidigt som de skall kunna operera i oskyddade miljöer, bidrar detta till stora säkerhetsutmaningar.

Dagens IoT är i stort behov av standardisering och klara strukturer för hur tekniken ska implementeras samt samverka med varandra på ett säkert sätt. Utmaningarna ligger i att säkra tekniken samt informationen som tekinken bidrar med. Denna rapport ger en introduktion till vad IoT är och hur det kan användas samt vilka hot som IoT kan möta i avseende till informationssäkerhet. Utöver detta så förser rapporten läsaren med förslag om hur man eventuellt kan lösa de fundamentala behoven av autentisering och säker kommunikation. Lösningarna som läggs fram är baserade på både nutida lösningar och teknik som är under utveckling inför framtiden. Nutida lösningar är baserade på säkerhetsprotokoll som IPsec och DTLS som används i en miljö som sträcker över internet och in i ett 6LoWPAN nätverk. Den autentiseringslösning som tagits fram grundar sig på PKI och förtroendemodeller för certifikathantering.

För framtida arbete presenteras flertalet vidare fördjupningsområden där denna rapport kan användas som grund. Dessa fördjupningsområden inkluderar vidare analys av sårbarheter och implementation av de lösningar som tagits fram.

Nyckelord:

Internet of Things, IoT, informationssäkerhet, identifiering, autentisering, säker kommunikation

Acknowledgments

We would like to thank:

Professor Gerald Q. Maguire Jr. for being our academic advisor and for his contribution by defining a starting point for this project and providing valuable input.

A special thanks to Emma Andersdotter at Combitech AB for offering and supervising this Bachelor's thesis project and participating in our interview study. By providing guidance Emma helped us define a purpose and a well-structured work process together with many valuable discussions.

Johan Thulin at Combitech AB for offering and supervising this Bachelor's thesis project and being a part of our interview study.

Additional people we would like to thank;

- Patric Brännström – for valuable discussions and tips
- Anonymous Combitech AB employee – participant in interview study
- Anonymous Combitech AB employee – participant in interview study
- All other employees at Combitech AB who were involved in this project.

Stockholm, June 2015
Cristian Bude and Andreas Kervefors Bergstrand

Table of contents

List of Figures

List of Tables

List of acronyms and abbreviations

AH	Authentication Header
AI	Artificial Intelligence
CA	Certificate Authority
CASAGRAS	Coordination and Support Action for Global RFID-related Activities and Standardisation)
CoAP	Constrained Application Protocol
DTLS	Datagram Transport Layer Security
ESP	Encapsulating Security Payload
HIS	Hardware Intrinsic Security
HSM	Hardware Security Module
IC	Integrated Circuit
ICT	Information and Communication Technology
IERC	IoT European Research Cluster
IKE	Internet Key Exchange
IoE	Internet of Everything
IoT	Internet of Things
IPsec	Internet Protocol Security
ITU-T	International Telecommunication Union ITU Telecommunication Standardization Sector
KINK	Kerberized Internet Negotiation of Keys
MAC	message authentication code
NIST	(United States) National Institute of Standards and Technology
PKI	Public Key Infrastructure
PUF	Physically Unclonable Function
RFC	Request for Comments
RFID	Radio Frequency Identification
ROI	Return on investment
SA	Security Association
SAB	Security Association Database
SEND	SEcure Neighbor Discovery
6LoWPAN	IPv6 over Low power Wireless Personal Area Networks
SPD	Security Policy Database
SPI	Security Parameter Index
SRAM	Static Random Access Memory
TLS	Transport Layer Security
2FA	Two-factor Authentication
UN	United Nation
URL	Uniform Resource Locator
WWW	World Wide Web
TCP	Transmission Control Protocol

UDP	User Datagram Protocol

1 Introduction

This bachelor's thesis project was conducted during Spring 2015 by two KTH Royal Institute of Technology students at Combitech AB. This chapter contains a comprehensive introduction to the Internet of Things (IoT). Following this introduction, Chapter 2 provides extensive information about security within IoT.

1.1 Introduction to the Internet of Things

The Internet of Things (IoT) is a new, but at the same time an old term. It was already mentioned by Kevin Ashton in 1999, while holding a presentation at Proctor & Gamble. He used the term to link the idea of radio frequency identification (RFID) to the then new topic Internet [1]. Since then the use of this term has blossomed and major companies have predicted an increase in IoT [2, 3, 4]. One prediction is that the number of connected things in the world will have a thirtyfold increase between 2009 and 2020, thus by 2020 there will be 26 billion things that are connected to the Internet [2]. The reason IoT has become so huge depends partly on two things: Moore's law and Koomey's law. Moore's law states that the number of transistors on a chip doubles approximately every two years [5]. This has enabled people to develop more powerful computers on the same sized chip. Intel, a well-known semiconductor chip maker had during 1971, 2300 transistors on a processor and by 2012 their current processors contained 1.4 billion transistors [6]. This is an increase of approximately 610 000 % and it is expect that this trend will continue.

Koomey's law explains that the number of computations per kilowatt-hour roughly doubles every one and a half years [7]. Kevin Ashton states that these two laws have together enabled us to create powerful and energy efficient computers. By turning the graph for Moore's law upside down it can be interpreted as the size of a computer (of a fixed capacity) is halved every two years. Doing the same thing to Koomey's law can be interpreted as the amount of energy needed to perform a computation is dropping at a rapid rate [8]. Combining these interpretations tells us that we can perform the same amount of computations on increasingly a smaller chip, while consuming decreasing amounts of energy - hence computations are becoming more energy efficient. The potential result is a small, powerful, and energy efficient computer which enables us to provide more advanced services using less chip area and at a lower energy that what has been possible before.

Defining the term IoT can be somewhat difficult because it has many definitions depending on who is defining the term [9]. The basic concept of IoT is to connect things together, thus enabling these "things" to communicate with each other and enabling people to communicate with them [10]. What these things are varies depending on which context the term is used and the aim of using the thing. In this thesis we have chosen to follow the definition of IoT proposed by ITU's Telecommunication Standardization Sector (a United Nations agency which specializes in ICT): "*… a global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies*". Interconnecting the physical world with the virtual world and applying this concept to all things opens up new possibilities in the sense of being able to at any time access anything from any place. Providing new possibilities will also generate new threats, security risks, and expose vulnerabilities in the unexplored world of interconnected everything. "Things" in the physical world are objects that physically exist and from the perspective of IoT we are able to sense, operate, and connect to these things, while in the virtual world "things" are objects that can be stored, accessed, and processed [11].

IoT involves sensors in order to collect information. Sensors are already being used in daily life, however most people may not realise it. Smartphones contain different kind of sensors, such as accelerometers, cameras, and GPS receivers. Built-in sensors are nothing new in today's society. Kevin Ashton said that IoT is already happening, but we might not see it compared to Smartphones which can both be seen and touched. RFID is such an IoT-technology that exists but is not necessarily seen; so the development of IoT might progress a long way before it is visible for everyone [8].

1.2 Background for the Internet of Things

The most vital part of achieving IoT is communication, because in order to interconnect different devices they must be able to communicate. All other properties, such as sensing, manoeuvring, being able to capture, store, and process data are unnecessary; unless your device specifically requires one of these properties. However, the ability to communicate is essential when labelling a device as an IoT device. How this communication is performed is less important, since the actual physical and link layer communication within IoT can be realized in many ways.

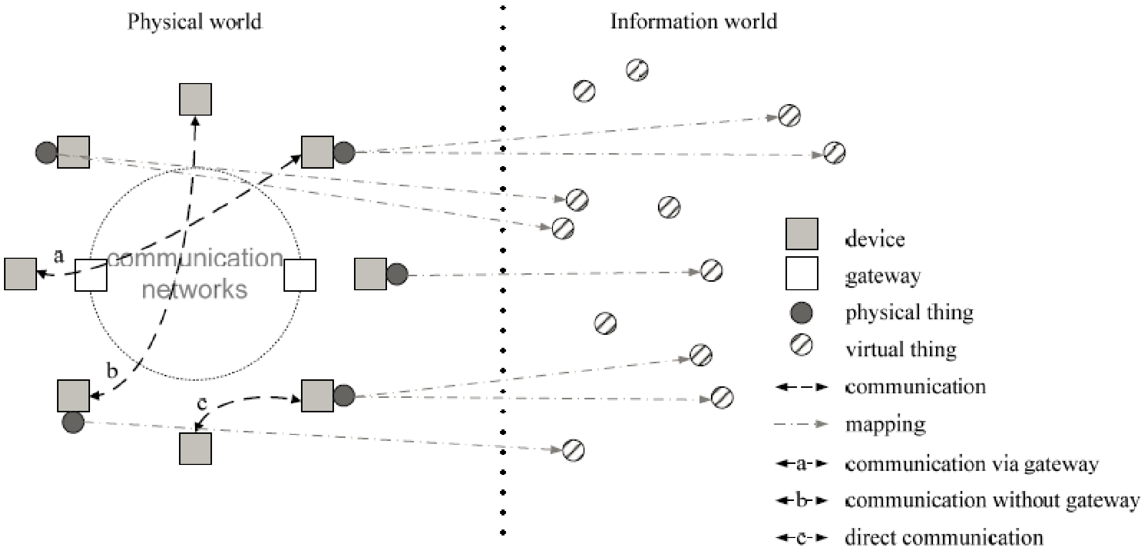


Figure 1—1: Overview of the Internet of Things (Used with permission from the author(s) of [11]).

Case C in Figure 1—1 shows that devices are not always required to communicate through a communication network. For example, if two devices are close to each other it might be simpler to directly communicate via for example radio using technologies such as Bluetooth or ZigBee (protocols which both enable direct communication). In contrast, in Case A in Figure 1—1 a device might communicate via a gateway using one protocol (such a IPv6 over Low power Wireless Personal Area Networks (6LoWPAN)) and then the gateway could communicate using another protocol (e.g. IPv4) over a communication network such as the Internet. Case B in Figure 1—1 illustrates two devices which are directly communicating with one another without requiring a gateway where both devices are directly connected to the communication network and thus are able to communicate even if they are located in different places.

A physical thing can be mapped into the information world via one or more virtual things, while virtual things do not necessarily need to be associated with any physical thing and can exist independently of any physical existence. For example, a physical thing might execute multiple applications and thereby have multiple identities in the virtual world. Similarly a virtual thing might also have many identities in the virtual world. For example, a virtual thing could be a video (file) on a USB-drive. Such a file might have multiple file names that refer to it and it might even have multiple instances (copies), potentially these “copies” might have different encodings, resolutions, etc.

How does one differentiate an IoT device from any other device? Table 1—1 states some fundamental characteristics for IoT. These characteristics may provide a clearer picture of the actual differences between IoTs and other devices [11] .

Table 1—1: Characteristics of the Internet of Things

Characteristics	Description
Interconnectivity	Everything can be connected to the global information and communication infrastructure
Things-related services	Provides things-related services within the constraints of things, such as privacy and semantic consistency between physical and virtual thing.
Heterogeneity	Devices within IoT have different hardware and use different networks but they can still interact with other devices through different networks. (i.e., Case A in Figure 1—1. using different protocols or hardware, but still be able to communicate)
Dynamic changes	The state of a device can change dynamically, thus the number of devices can vary. (Device states: connected, disconnected, waking up, and sleeping)
Enormous scale	The number of devices operating and communicating will be larger than the number of devices in the current Internet. Most of this communication will be device to device instead of human to device.

Interconnectivity is the basic characteristic for IoT since the whole concept is built upon the idea of being able to interconnect everything (despite the traffic going through different networks). Things related services resolves around devices being constrained by its CPU performance, memory, and power which limits what a device can do, when it can do it, and how often it can do it.

To provide semantic consistency a physical thing reporting temperatures at some intervals may be mapped to a virtual thing that tries to estimate the temperature between measurements and thus may report a different temperate value than the physical value. When the next measurement arrives the virtual device may or may not update its estimate in order to maintain consistency with the physical thing.

In Table 1—1 the biggest challenge will be supporting heterogeneity because there are a lot of different protocols in use. Interacting with multiple devices through multiple networks will be challenging from both security and technical perspectives, because the protocols may differ depending upon whether the device is communicating through one interface or another (e.g., wide area cellular radio, Ethernet, or Wi-Fi). Therefore, there are some requirements relevant for IoT, such as security and privacy protection. If everything is connected, then multiple security threats will arise causing confidentiality, integrity, availability, and authenticity to become more important – especially because there will be more data and services available and because more and more activities will depend upon this information. Security also includes privacy consideration, since data collected by for instance a sensor might contain information that is sensitive personal information. Integrity has to be considered in all stages (sensing, storing, transmission, etc.) that means that the security within IoT will have to adapt to a variety of devices and networks [11].

A thing that reports a geographical location can for privacy reasons add noise to its position (i.e. degrade its accuracy) thus the physical location compared to the virtual location can differ. This prevents the device from having an exact location mapped to it thus protecting spatial privacy.

1.2.1 The IoT reference model

The ITU-T has defined a reference model for IoT. This model is divided into the four layers: application layer, service support and application support layer, network layer and device layer (see Figure 1—2). Each one of these layers also includes management and security capabilities. As shown in the figure these capabilities have both generic and specific capabilities that can cut across multiple layers.

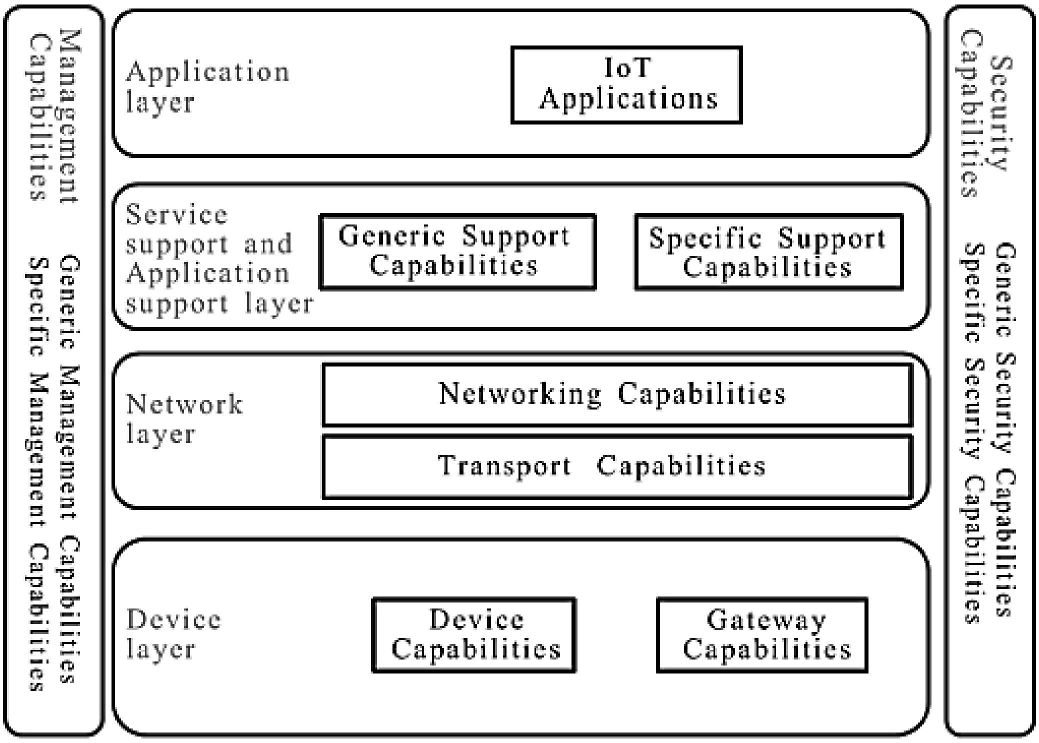

Management Capabilities
Generic Management Capabilities
Specific Management Capabilities
Application layer
IoT Applications
Service support and Application support layer
Generic Support Capabilities
Specific Support Capabilities
Network layer
Networking Capabilities
Transport Capabilities
Device layer
Device Capabilities
Gateway Capabilities
Security Capabilities
Generic Security Capabilities
Specific Security Capabilities

Figure 1—2: ITU-T reference model for IoT. Taken from Recommendation ITU-T Y.2060 and used with permission from author(s).

The application layer contains IoT applications which require certain support capabilities from the underlying layer to function. The service and application support layer consists of generic support capabilities which can be used by IoT applications, examples of such capabilities could be data processing or storage. The specific support capabilities are those other than the generic capabilities which are required to create support for diversified applications [11].

The network layer is divided into networking and transport capabilities. The networking capabilities provide relevant control functions for network connectivity, while the transport capabilities focus on the transport of IoT service and application specific data. At the bottom of the model, there is the device layer in which the device capabilities include direct and indirect interaction with the communication network. Unlike direct interaction, indirect interaction requires a gateway to be able to send and receive information via the network. Two other capabilities are *ad hoc* networking and sleeping and waking up which enable devices to connect in an *ad hoc* manner and saving energy (respectively) [11].

The device layer also includes gateway capabilities to support devices connected via different types of wired and wireless technologies by supporting multiple interfaces. In some situations, protocol conversion is needed to support communication between devices using different protocols at the device and network layer [11].

Generic management capabilities include device management (such as remote device activation, de-activation, diagnostics, and firmware or software updates) and local network topology, traffic, and congestion management [11].

The generic security capabilities are independent of the application and include authorization and authentication at the application, network, and device layer. Moreover, all of the layers have their own individual capabilities. These include:

At the application layer	application data confidentiality and integrity protection, privacy protection, security audit and anti-virus;
At the network layer	signalling data confidentiality and integrity protection; and
At the device layer	device integrity validation, access control, data confidentiality, and integrity protection.

Both the specific management and security capabilities are closely coupled with application-specific requirements, for example mobile payment [11].

1.2.2 How is the term IoT used today?

Since 1999 the term IoT has been used in many places and in many ways. Multiple research papers, books, and white papers about IoT have been written in order to help both the public and companies understand what IoT is. Many definitions of IoT have been independently introduced by both individuals and companies [9].

Technical companies that are already somewhat involved in IoT and who believe that IoT has a business potential for their future mostly use the term to describe a way of improving efficiency of production and innovation. Cisco defines IoT as concept where more and more things will be connected to the Internet in order to ease people’s daily life. However, as we connect more things, the need for IPv6, big data, and cloud computing will increase and the concept of IoT will transition into an Internet of Everything (IoE). Cisco views IoT as a phase where the number of connected devices increases, while this phase changes once everything connected [12].

IBM has a definition of IoT which is more about connecting systems together, rather than just connecting devices together; thus, their focus is on creating a system of systems. They describe IoT as a means to create a smarter planet. They split these means into two parts: “One is to be more efficient, be less destructive, to connect different aspects of life which do affect each other in more conscious, deliberate and intelligent ways. But the other is also to generate fundamentally new insights, new activity, new forms of social relations” [13].

Individual definitions include that given by Dr John Barrett, Head of Academic Studies for Embedded Systems Research at Cork Institute of Technology in a TEDx talk on the requirement for IoT: In the context of IoT all things will need a unique identity (IPv6), ability to communicate, in some way sense (see, smell, touch, etc.) and to be controlled. With all the collected data there is a need for a practical and efficient way to present the data that is relevant in a certain context. Deciding what is relevant becomes a core question. It is up to the things themselves to decide what is relevant and what is not. In some cases the “relevant” data may be misused in a way that negatively effects people. For example, a device monitoring your health can be used to notify the hospital if your health is in critical condition. However, by using the same information as the hospital, your insurance company automatically *increases* your health insurance premium by 25% [14].

Another interesting individual definition is given by Kevin Ashton, who continues to give presentations regarding IoT. Like many others he sees IoT as fundamental for creating solutions for future problems. He defines IoT as computers sensing the real world by themselves and for themselves, thus information about things in the world can be available via the Internet. The problem with IoT is not deploying sensors everywhere; but rather the creation of systems that are able to exploit all of the available data and automatically figure out what it means [8].

CASAGRAS (Coordination and Support Action for Global RFID-related Activities and Standardisation) is a research project funded by the 7th European Framework Programme. The project focuses on international dimensions relating regulations, standardisation, and other requirements for realizing IoT. CASAGRAS defines IoT as following, *"A global network infrastructure, linking physical and virtual objects through the exploitation of data capture and communication capabilities. This infrastructure includes existing and evolving Internet and network developments. It will offer specific object-identification, sensor and connection capability as the basis for the development of independent cooperative services and applications. These will be characterised by a high degree of autonomous data capture, event transfer, network connectivity and interoperability"*. Their analysis concludes that the development of IoT will require attention to fundamental features, infrastructure, architecture, and technical significance. Initially IoT will require a basic framework in order to define and accommodate the development of IoT. The use of this framework is not meant to remove the need for a defined goal [15].

As mentioned in the introduction, the definition by ITU-T highlighted the enabling of services and the interconnecting of things. This should be made possible by using existing and evolving communication technologies. ITU-T defines a three dimensional space in which IoT adds one of the dimensions (any**thing** communication) to the information and communication technologies which already provide the two others: "any **time**" and "any **place**"). In other words, previously we could communicate at any time and place, but with IoT we can communicate with any "thing".

Even though businesses, individuals, and papers explain IoT in slightly different ways, the similarity of their definitions centres on the interconnecting of things. The difference in their definitions is how they present the concept. Businesses mostly focus on the possibilities within IoT with regards to *efficiency* and *innovation*, but do not mention the security threats which may arise. This does not mean that these businesses are unaware of potential risks and that they do not have a suitable plan regarding IoT (although this could be true). However, business may simply choose *not* to publicly announce the risks they see with the concept of IoT nor how they plan to secure it. For a business it is always valuable to possess information which your opponent does not. While at the same time security via obscurity has been found time and again to not really provide security!

The research papers the focus is often on defining one or more standards for IoT. Today, there is some common ground between individuals and research papers when writing about IoT as both highlight the possibilities of IoT and emphasize the need for privacy *and* security.

1.2.3 Where is the term IoT being used?

The term IoT is being used in different contexts, such as the body, homes, cities, industry, and the global environment.

- In terms of the body, IoT enables sensing and connectivity, for example tracking activity, health status, and other relevant information *could* improve not only the user's daily life, but also their future health by preventing bad habits. However, this could come at the cost of a tremendous decrease in personal integrity and personal autonomy. Hence there are both individual and societal issues that have to be addressed with this sort of IoT.

- When talking about the home, IoT is often considered in terms of remote and local monitoring and management of different home electronics and lights, or simply to keep plants in the yard alive by using an automatic watering system. Today this is becoming a very important area as more and more areas are facing shortage of water, hence traditional approaches to watering house plants and gardens are no longer feasible.
- In correlation to cities, the term IoT is used to describe systems that effectively gather and process information generated by various infrastructures, for example monitoring centres for traffic lights, street lights, camera surveillance and the power grid. These systems offer the potential to improve the flow of vehicles and people through the city centres and also greatly improving the energy efficiency of transport systems, while also improving personal and societal safety.
- Optimizations of operations, boosting productivity, saving resources, and reducing costs are typically the main goals of IoT solutions applied in industry. For example, industry might use IoT to keep track of business assets, improve environmental safety, and maintain quality and consistency in a production process. This is not only a matter of companies seeking to be “green” but also because there are very substantial economic advantages to understanding how to do better process control (in terms of maintaining quality), but also lessening the harmful effects upon the environment.
- Last, but not least important, is environmental monitoring where IoT can help us understand and better manage those resources we have. Sensors can help protect wildlife, track water usage and flows, monitor local weather, monitor use of natural resources, or give warnings before and after natural disasters to prepare people for what is to come [16]. In fact, it appears that to achieve high environmental efficiency requires increasing use of information technology (whether this is in production, consumption, recycling, or post-recycling phases).

1.2.4 Conclusion: What is IoT?

IoT includes different *objects* with different *capabilities*, which have a common way of *communicating* (a communication chain through a communication network) for enabling transfer of information, where this information is understood by two or more *objects* in order to make a *process* more efficient; frequently by minimizing human factors and interaction.

Objects include both virtual and physical objects, but are not limited to:

- Electronic devices (e.g. computers, mobile phones, televisions, machines, and robots) and
- Sensors (connected through devices or gateways)

Communicating includes:

- Different protocols and technologies for sending digital or analogue signals through nodes (e.g. Constrained Application Protocol, File Transfer Protocol, Hypertext Transfer Protocol, etc. in Local Area Networks, Wide Area Networks, Body Area Networks, Wi-Fi, Ethernet, fibre optic links, radio etc.)

Capabilities include, but are not limited to:

- Gathering information,
- Processing information,

- Storing information, and
- Presenting information.

A process could include:

- Tracking health information,
- Heating your home,
- Lighting public streets, and
- Keeping track of assets.

An example of non-IoT is a single object speaking its own language (even with the use of protocols) and potentially connected to a communication network (e.g. Internet), but no other object is able to interpret this data and therefore no other object can contributing with any functionality or usefulness to this non-IoT device. However, as soon as there is something on the other end of the communication path that can use the same protocols, then it is possible to establish communication and potentially increase efficiency.

A practical example of IoT is the Bigbelly smart waste and recycling system, shown in Figure 1—3. In this system stations (*objects*) made for waste collection, monitor (*capability*) and report (*communicate*) station fullness and station-specific data remotely, in this case to the Bigbelly cloud (*object*). This helps garbage collectors know when and where a station needs emptying (*process*) which historically has been a guessing game [17].

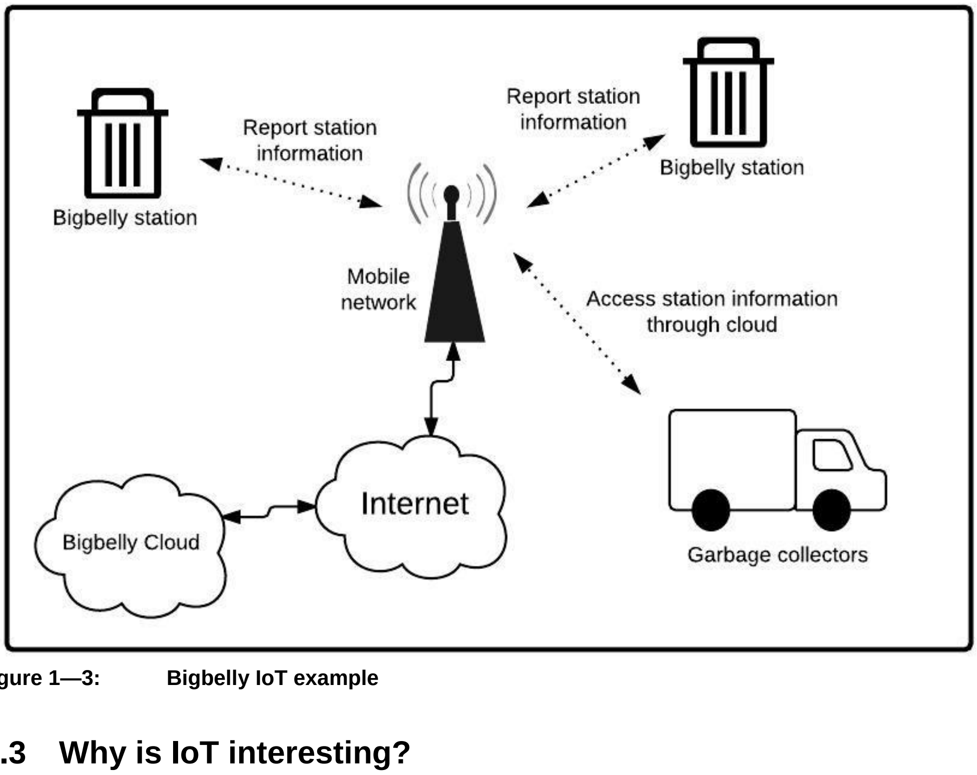

Report station information
Report station information
Bigbelly station
Bigbelly station
Mobile network
Access station information through cloud
Internet
Bigbelly Cloud
Garbage collectors

Figure 1—3: Bigbelly IoT example

1.3 Why is IoT interesting?

Together with the expansion (in numbers of people) and the goal of a sustainable society, we need better ways to collect and distribute information (generally over the Internet), while maintaining accuracy, reliability, relevancy, and security. Many years ago, almost every piece of digital information was typed, recorded, or in some other way created by human beings. Humans are fundamentally limited in the rate at which they can generate information. However, computers and other devices can generate information without any human interaction, which increases the possibility to collect sufficient information to reduce unnecessary loss and costs. For example, by monitoring the vibrations of a motor, we can estimate when we should repair or replace the motor, while avoiding the need for constant attention or periodic check-ups [18]. Additionally, we can schedule when this repair or replacement is done, thus increasing the effective performance of the vehicle, escalator, or other device that the motor is powering – while avoiding the need to perform emergency repairs and avoiding being inoperable when there is the greatest demand.

However, in order to take appropriate actions based on our decisions, we need to know that the information that we are basing our decisions on is accurate, reliable, and correct; in other words, that the information exchanged between the things and ourselves is secure and accurate.

In order to achieve accuracy, reliability, and relevancy in the enormous amounts of generated and processed data, there is need of transferring the human intelligence and appropriate security mechanisms to the systems in use. Artificial intelligence (AI) is the word used to describe computer systems with intelligent behaviour, behaviour such as representation, searching, reasoning and learning, which are the four fundamentals of AI.

A system with AI needs an internal *representation* of a problem or related knowledge to be able to know when a problem arises. If we reconnect to the practical example of Bigbelly (see Figure 1—3) the stations need to know what possible problems it may encounter. From the garbage collectors perspective (and thereby the station), one problem is when a station becomes “full”.

After a problem has been identified the next step is to find out what to do when the station becomes full, which is often done by using different kinds of *search* methods. When relevant information associated to the problem has been found, *reasoning* together with the knowledge is used to find a fitting solution. Logically, this would be to tell the garbage collectors to empty the station.

Most systems with AI also have the ability adapt and optimize if necessary, which is done by *learning* based on historical statistics for example. This is used in the Bigbelly example to reduce collection frequency, optimize routes and reallocate resources to other tasks [19].

1.3.1 Where does the intelligence lie?

The activities of Artificial Intelligence (AI) can be spread around in IoT and does not require that all the activities occur in the same place. In IoT, the collection of all these activities is what creates the AI. The flow of data in regards to sensing and processing can be presented in different ways. The flow can be as simple as an object acquiring data through a sensor, which it then processes and finally transmits in the form of a data packet, as shown in Figure 1—4 [20].

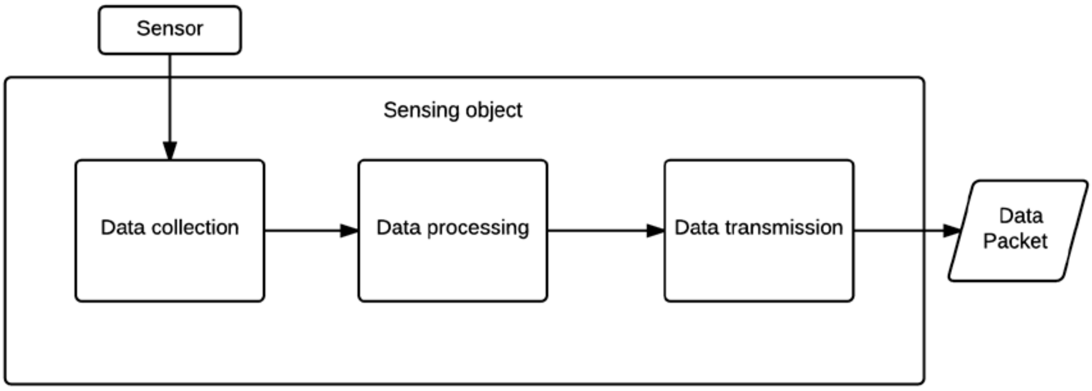


Figure 1—4: Example of dataflow

Another example is a framework for an industrial park which has a system that is capable of perceiving, analysing, and predicting future events. The scenario is an enterprise where all its power equipment is controlled within the IoT. The system will be able to predict different events, such as if the power system will reach its peak by the next measurement and through calculation be able to predict if the peak will exceed the expected power limit. Using these predictions the system can affect these predictions by affecting the power usage by for example, lowering the power of electric equipment, shutting off electronic equipment, or utilizing alternative power sources before the predicted over limit occurs [21]. Figure 1—5 illustrates a scenario where the equipment senses its power usage and collects data, transfers it to a transmission platform, which in turn collects data from multiple objects (i.e. all the equipment in a building) and then sends the collected data to a remote third party service provider who processes all the data and takes some action that in the end affects the power usage within the enterprise.

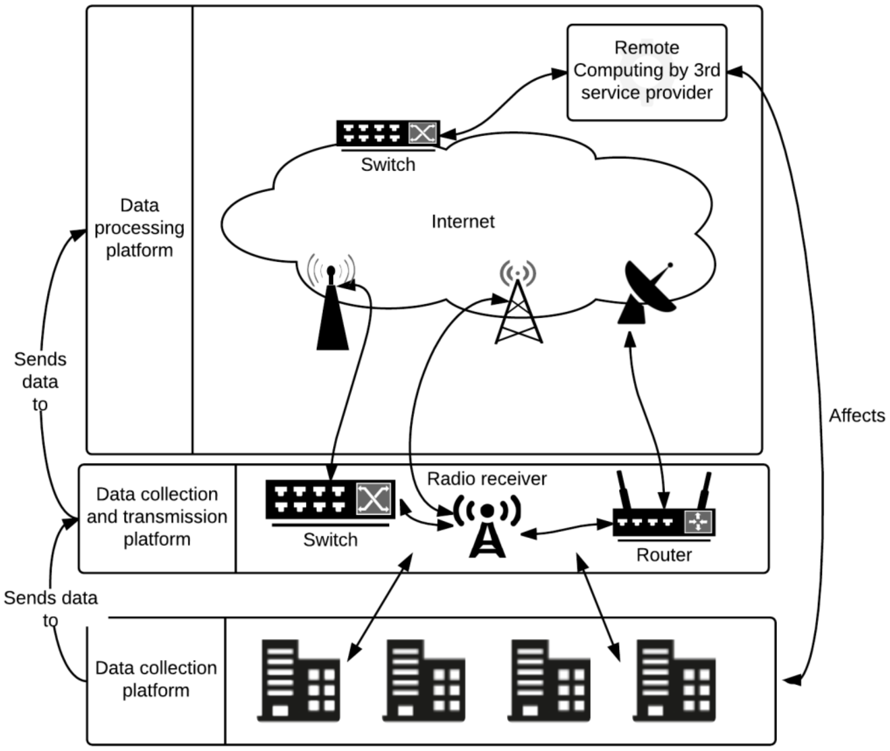


Figure 1—5: Framework of an IoT industrial park

The simplest way of describing a data processing flow is a monitoring object collecting data which is sent to a computation device that processes and analyses the data. The computation device then sends the result to a terminal which executes a command based on the result or simply presents the data to for example a user [22]. This data processing flow can be seen in Figure 1—6.

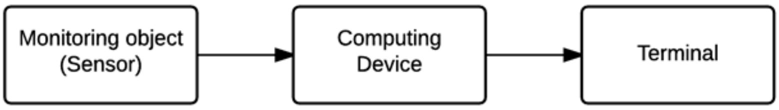


Figure 1—6: Simple data processing flow

Nothing states that within IoT the processing of data occurs only once during the dataflow in IoT. It is possible to combine the flow in Figure 1—4 and Figure 1—6. The sensor itself might do some processing before the data is sent to a larger collection point which in turn sends the data to a processing point. The use of remote processing is especially relevant when the system consists of multiple objects that together provide the data necessary in order to decide if for example a command needs to be executed or not, as in Figure 1—5. The artificial intelligence (AI) is not necessarily positioned in the same place (platform/device) since their placement will depend upon the structure of the IoT-environment.

In the case of the Bigbelly system, the garbage bins are equipped with an integrated circuit with a processor that monitors the garbage bin, thus realizing a fully automated system which senses trash level, fullness, and machine status. Here some of the computation is done by the object (garbage bin) itself and the result sent to the terminal, which in this case is the Bigbelly cloud. The Bigbelly cloud analyses all the data it receives from the different garbage bins and presents this data to the user in different ways, such as a map the location of these bins and their status (trash level/fullness). In this case the terminal only presents the results to a user and does not execute a command based on these results [17].

1.4 Future of the Internet of Things

When connected to the Internet, the possibilities that others can see us, hear us, and control devices is greatly expanded through the deployment of IoT. Moreover, privacy and personal integrity concerns arise as more data is collected about our activities and personal information, in the form of locations, habits, or financial account numbers [23]. Additionally, many of the decisions and actions that will be taken based upon and using IoT will have real-world costs, risks, and benefits. For these reasons, we need to ensure that security considerations are part of the design process and not something that is added late in the development of each IoT device.

When deploying things that are capable of connecting to the Internet it will be important that implementations are done correctly, otherwise systems and their information might be exposed to attacks. BMW, a major vendor in the car industry, recently had an Internet related security-hole discovered. The problem was caused by careless implementation. The fault was an optional car feature called ConnectedDrive which connects to the Internet via the public cellular network using a SIM card. The feature allows the owner to remotely switch on the heating or air conditioning, sound the horn, and lock or unlock the car using their smartphone [24]. The problem lay in the car's communication which was unencrypted, this enabled people, other than the owner, to open the locked car and left 2.2 million cars exposed. Luckily a German automobile club called ADAC (Allgemeiner Deutscher Automobil-Club) discovered this and notified BMW before any criminal offenses caused by this problem were reported. The solution was to switch on encryption for the communication and all cars were fixed by 31 January 2015 [25].

Even though the source of the problem was small and the solution was simple, this example shows how one small mistake can be both extensive and expensive. It is for this reason that security considerations *must* be a part of the design process. Unfortunately, companies may not have sufficient time and/or economic resources to perform sufficient tests before deploying their product, thus trading future risk and costs against current costs. Today's rapid development is also a factor because it forces companies to keep up with the market demands.

Security wise, encrypting communication might not be relevant in all cases. This depends on which elements of information security are important in each specific case. If integrity, availability, and authenticity are important, then encryption is a vital part of the communication process. In contrast, a weather sensor might not need to encrypt its traffic, but might only need to add a cryptographic hash to ensure integrity and authenticity. As a business you might not be concerned if queries and answers to/from the sensor are visible on the Internet, as long as the integrity of the sensor's values and their authenticity are ensured. Security within IoT will depend on which threats one wishes to protect themselves against. In some cases, it will also depend upon explicit decisions to make data open so that it can be used by others, thus facilitating new and unexpected applications.

1.5 Problem definition

Since IoT is a relatively new concept, it is still largely unknown and unexplored by many companies and employees in industry. This limited knowledge may cause them to be afraid of, or as in the previous example, totally unaware of the potential security and privacy issues connected to their deployment of IoT [26]. This is why many businesses want to know more about the potential threats, benefits, disadvantages, and solutions regarding security in conjunction with IoT. Additionally, they need to know what competence in information security is necessary in order to realize *cost effective* security in conjunction with their deployment of IoT. This knowledge and competence should help facilitate their transition from a non-IoT-business to an IoT-business, as it will enable both employees and management to understand & address their doubts & concerns in terms of their investments and the resulting security risks. In this way, managers can make a balanced risk-benefit analysis of the adoption of IoT for a specific application or family of applications.

1.5.1 The “Things”

The concept IoT includes all kinds of different technologies and every possible way to communicate between (virtual or physical) objects via the Internet. The breadth of this concept makes it rather complex because of the heterogeneity of components. Since every type of device may use its own specific hardware and software, there is a wide range of operating systems and applications that have to be considered. In some cases, the device may not even have an operating system, for example, there are devices that only have a network interface, a driver, and an application generating (or sinking) data.

Security threats can emerge from any of the layers shown in the IoT reference model, see Figure 1—2ö. The sources of threats include authenticated and in most cases, non-authenticated users that have access to any of these layers. With all the different combinations of hardware and software it is very hard to define a common security model, let alone a single globally applicable model.

1.5.2 Communication

Devices from different vendors often speak different protocols. In some cases these protocols are proprietary, hence unknown to public. Experience has taught us that secure protocols demand open peer review to provide robust assessment and thus attract wide acceptance and use [27].

Today, IoT devices increasingly have one thing in common due to their use of Internet Protocol (IP) at the network layer in the protocol stack. The reason that these devices increasingly use IP is to enable communication through the Internet. The famous hourglass model (see the left-hand side of Figure 1—7) shows the concept pretty clearly, with everything on top of IP and IP on top of everything [28]. Note that the right-hand side of the figure shows an alternative approach where chunks of data are named and these named chunks are requested.

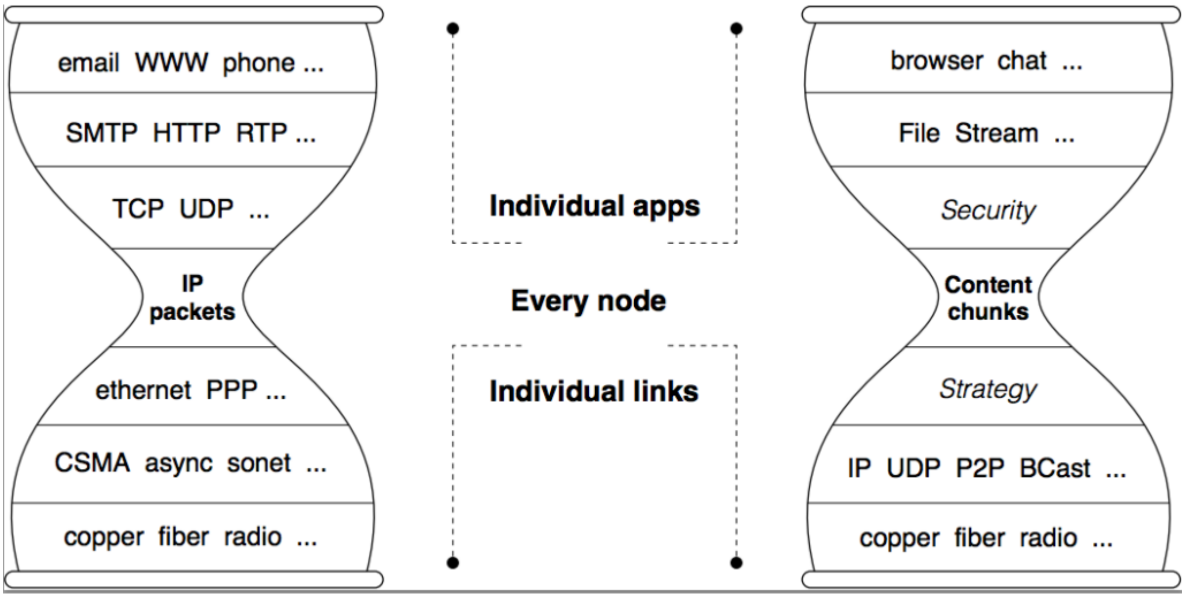


Figure 1—7: **IP stack with IP in the middle compared with content based addressing with content chunks in the middle. (Taken from Named Data Networking under the Creative Commons License 3.0. [94])**

This has lead in recent years to an IP based network of "things". Whether the thing is a sensor or actuator, it increasingly utilizes IP based communication to communicate with a controller. This means that a single sensor or actuator can be seen as part of IoT, as long as it is connected to a device with access to the Internet. Other protocols can and will be used, depending upon what other network attached devices can do with the information or the device; however, as long as the device utilizes IP & is connected to the Internet or can talk to a proxy that utilizes IP & is connected to the Internet, then it should be possible to remotely communicate with the device. The main problem will be what to communicate and what processing has to take place to achieve the desired objective.

The current or at least the most widely used version of IP today is the Internet Protocol version 4 (IPv4). However, there is a practical problem with utilizing IPv4 in the context of IoT as there are insufficient addresses available to directly address each of the things that are likely to be part of IoT. This limitation in the number of addresses occurs because the address fields are only 32-bit long. As of today in most places in the world, we have effectively already run out of addresses [29].

The latest version of IP (IPv6) [30] has been deployed in many systems. IPv6 uses 128-bit long address fields, which gives us a total of 2^{128} unique addresses which is 7.9 octillion times more than the number of IPv4 addresses. This size of address field basically gives us an unlimited number of addresses. For this reason adoption of IPv6 is needed for the rapid expansion in the number of devices that wants to communicate via the Internet. Moreover, essentially all modern operating systems and IP stacks support IPv6 and IPv6 includes support for IP security (see Section 2.5) [30].

In summary, the problem lies in how to identify, authenticate, communicate, and transfer data; while at the same time protecting individual's and company's privacy & integrity by securing the "things" and their data used in the process.

1.6 Purpose

The purpose of this thesis project is to provide both ourselves and our employer with basic knowledge about IoT. We started this process by asking and answering questions about IoT. Some of these questions have been raised in the previous sections, while others will be ask and answered in the following chapters. Some of the questions that we hope to answer are:

- What is IoT?
- How is it used today?
    - What does the market expect in terms of security?
    - How mature is the security of IoT security?
    - Which technical methods are currently used to realize the security of IoT? Which technical methods can be used in the future?
- Why use IoT?
- What does the future of IoT look like?
    - What are the characteristics of near terms problems and solutions?
    - What predictions can we make about longer term problems and solutions (Problems/Solutions which are difficult to solve with today's technology)?

The main focus of the rest of this thesis will be on the security of IoT, but in some cases we will look at the correlation between security mechanisms and their power consumption and cost. The later aspects are important as it is too easy to decide upon mechanism which will not be practical for real-world deployments. We will also try to identify what security problems have not yet been taken into consideration, together with solutions proposed for securing communication of IoT, secure storage, and authentication. We will focus on three aspects of information security (confidentiality, integrity, and availability) and analyse them both separately and as a whole. Note that this means that we are intentionally *excluding* considerations of non-repudiation; hence, this will remain for future work. Additionally, we will try to understand which of these are more or less important with regard to IoT. For example, some IoT devices will have short lives, for such devices, will availability beyond their expected lifetime be an important property or not?

It is interesting to see which security methods and solutions can be used in specific processes, for example when transferring data – communicating, saving data – storage, and accessing data - authentication in IoT. Relevant question s at issue could include:

- What threats are connected to a typical IoT environment?
- What is needed to secure a typical IoT environment?
- How can companies ensure that the communication between devices and the devices themselves is indeed controlled by the company itself?[*]
    - How do we communicate securely?
- How can mutual authentication be done in a secure way?
    - How do we identify an object?
    - How do we authenticate them?

Companies that want to know more about IoT and the potential security threats and solutions that exist today will greatly benefit from this thesis. After reading this thesis it should be clear to the reader what threats are most relevant and which potential security solutions within IoT exists today, especially when it comes to identification and authentication.

[*] Note that we do not consider the case of non-company controlled devices – since the focus of this thesis is on business users of IoT.

1.7 Goals

The goal of this degree project was to give the reader and Combitech AB a deeper insight into the Internet of Things. A deeper insight means:

- Trying to understand the concept IoT.
- Identify security challenges that need to be addressed in order to secure IoT-environments.
- Doing an internal interview study which will facilitate Combitech AB's future work in this area.

Based on the demands and goals of all parties, we strived to reach a common result which satisfies all sides (i.e. our examiner at KTH Royal Institute of Technology) [31], Combitech AB, and ourselves).

1.8 Research Methodology

The research methodology of this thesis project was based on both a qualitative interview study and quantitative literature study. Because of the limited time there was not room for a large number of interviews, therefore only a few interviews were conducted. From the literature study empirical evidence has been used to produce a partial solution to the question at issue. This solution was then applied to a problem for testing purposes.

The interview study was done as a request by Combitech AB and it provided us with qualitative information in the area of Information Security.

1.9 Delimitations

We limited our thesis to how to secure an IoT environment when establishing communication between a company server, devices, and gateways (see Figure 1—8). We looked at the first steps of securing an IoT-environment (i.e., identification, authentication, and sending data securely) and assess the security functions needed to counteract possible threats in the specific environment. Due to the limited duration of our thesis project, we did not take other protocols or hardware specifications into consideration (except that a device is more constrained than a gateway).

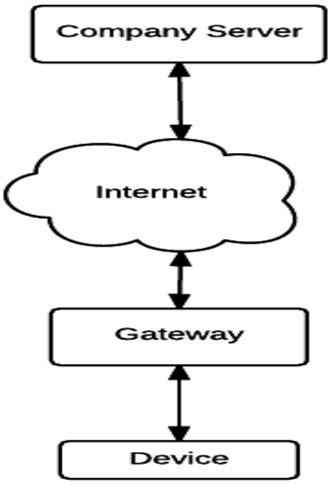


Figure 1—8: Delimitation of IoT for the purpose of this thesis project

1.10 Structure of this thesis

Chapter 2 presents relevant background information about Information Security and security thinking in conjunction with IoT. Chapter 3 presents the methodology and method when working with our thesis together with an explanation of the interview study. Chapter 4 present the results of our interview study and our proposed authentication method, together with our case scenarios that were used to analyse and counteracting different threats against our IoT-environment.

2 Background

This chapter provides basic background information about Information Security and how it is ensued in the Internet of Things. Additionally, this chapter describes threats associated with IoT-environments and what the most relevant characteristics are to secure them.

2.1 What is Information Security?

Information Security is an umbrella term for the processes and methodologies used to protect information, data and systems. In regard to Information Security, protecting mean preventing unauthorized access, use, disclosure, disruption, modification or destruction. Information Security has three key principles that can be taken into consideration. These are confidentiality, availability, integrity [32, 33, 34]. Accountability has become more important principle and is sometimes included among the three concepts by security companies (i.e., Combitech AB). These concepts are explained in Table 2—1.

Table 2—1: Principles used in Internet Security

Principles	Description
Confidentiality	Confidentiality is a concept which refers to the ability to protect data/information from people who are not authorized to view/access it.
Availability	Availability refers to the ability to ensure reliability and access to data/information when needed.
Integrity	Integrity refers to the ability to prevent unauthorized modification of data/information, thus assuring its accuracy and reliability.
Accountability	Accountability refers to the ability to trace modifications of information. A concept that's used to trace by who and when a change was made [33, 35].

The Parkerian hexad is an alternative framework for Information Security. Whilst keeping the principles of confidentiality, integrity and availability it adds the principles – utility, authenticity and possession [33, 34]. These terms are explained in Table 2—2.

Table 2—2: Principles unique for the Parkerian hexad

Principles	Description
Utility	Utility describes the usefulness of the information. Losing the encryption key for encrypted information renders the information useless. On the other hand, possessing encrypted information without matching key is also useless for an attacker.
Authenticity	Authenticity refers to ability to ensure authorship or claim of origin of information.
Possession	Possession is a more physical oriented concept and relates to losing possession of valuable information.

Losing possession does not necessarily mean that confidentiality is broken. Stealing valuable information which is encrypted relates to losing possession, but does not violate confidentiality since the thief cannot read the information. On the other hand, losing a valuable file can be disastrous if is the only copy.

How Information Security is approached can vary depending on the information that is to be protected. Information Security can be split into three questions in order to easier understand which concepts need to be taken into account in order to protect the information. The questions can be seen in Figure 2—1 and are as follows, “What needs to be protected?”, “What threats need to be addressed?” and “What needs to be assured?”.

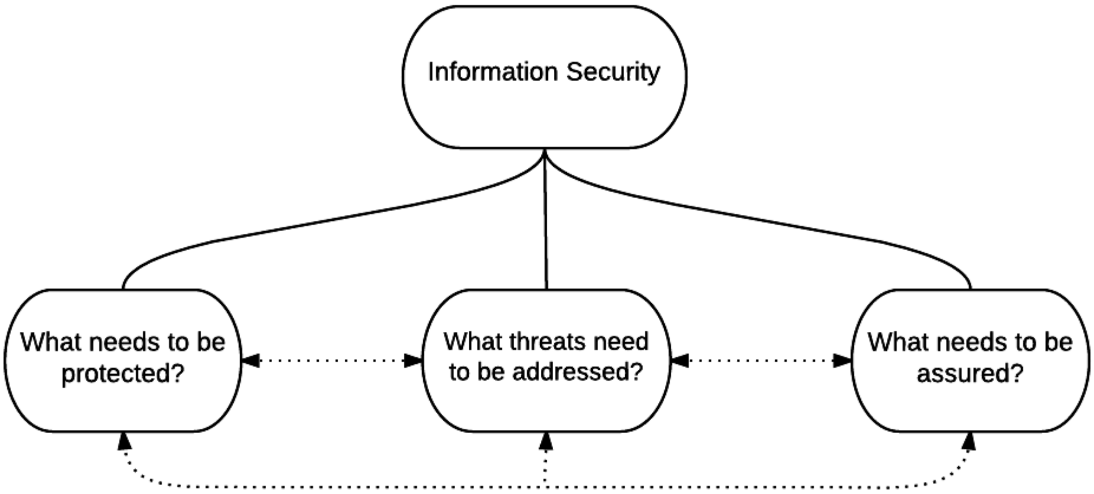

Information Security
What needs to be protected?
What threats need to be addressed?
What needs to be assured?

Figure 2—1: Questions for Information Security

2.1.1 What needs to be protected?

What is in need of protection? Is it a system or information? If the goal is to protect information it is relevant to identify if the information is confidential, internal or accessible to the public. If the information is public only availability, accountability and integrity are relevant. However, if the information is confidential the concept confidentiality also needs to be considered.

2.1.2 What needs to be assured?

Going back to system and information, this question relates to which of the concepts are desired to be met. Assuring that the system is unaffected by negative impacts means that availability is desired to be achieved.

There are many potential answers to this question and all answers (likely) apply one or more concepts. Examples are shown in Table 2—3 and all these answers relate to at least one concept.

Table 2—3: Example answers

Answers	Concept(s)
No unauthorized person can access the information	Confidentiality
The correct information is delivered	Integrity, Accountability
The receiver can confirm who the sender is.	Authenticity
The information is delivered	Accountability, Availability

2.1.3 What threats need to be addressed?

This question answers what kind of threats that are relevant for the system/information. Is a threat the destruction of the system or destruction/loss/falsification of information? These threats does not only apply to external threats but also internal. A threat is caused by an actor called threat agent, see Section 2.1.5 for the whole analysis chain. A threat agent can cause either deliberate or accidental threats, and different threat agents have different resources and probability to attack the system. A user with the lack of knowledge might destroy information by mistake. This mistake can also be made if the user is careless or intentionally destroys information.

2.1.4 Relations

These questions also relate to each other since they all answer the same kind of questions but from different angles. The questions what needs to be protected and what needs to be assured together contribute to finding an actual solution. The dotted lines between the questions in Figure 2—1 is a way of illustrating this.

2.1.5 Relationship between security components

When a system or environment undergoes an analysis the following method (see Figure 2—2) can be used. By starting with identifying a threat agent, the whole analysis chain will end up with a suitable safeguard. The safeguard will not eliminate the threat agent but it can prevent, or simply make it unsustainable to exploit the vulnerability that gives rise to the threat.

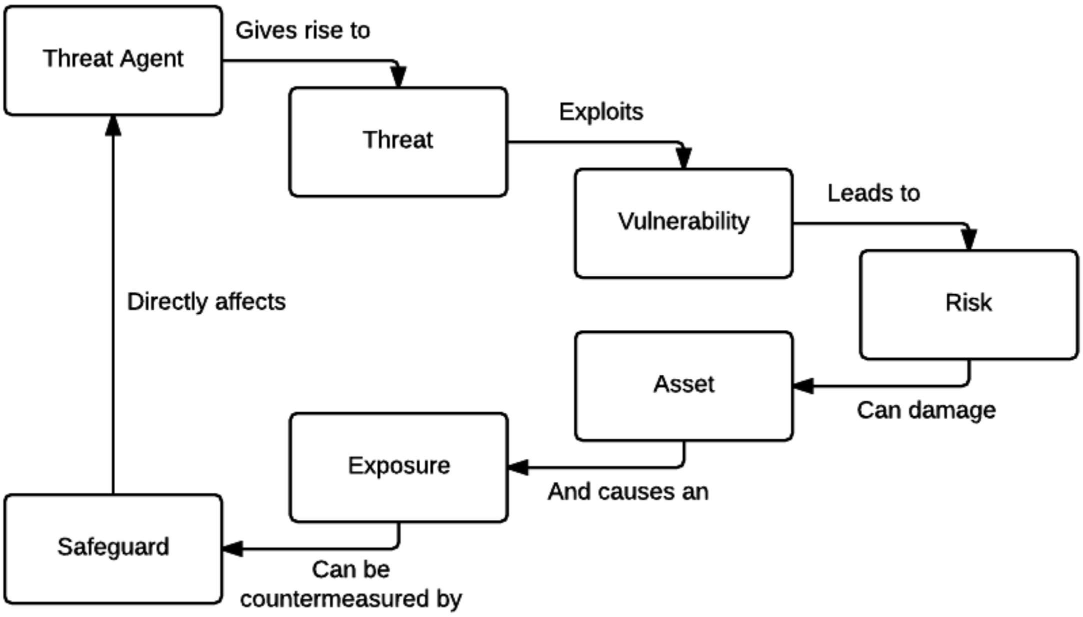


Figure 2—2: **Relationship between different security components [96]**

2.2 How is Information Security ensured in the Internet of Things?

The Jericho Forum is a series of publication guides from The Open Group that defines principles when planning for a de-perimeterised future, which fits very well to the concept of IoT. De-perimetarisation includes protecting an organisation's systems and data with a mixture of "secure" protocols, systems, and data-level authentication with the absence of a specific boundary between the organisation itself and the outside world [27]. In relation to IoT this describes a scenario when an organisation for example deploys weather sensors that collects information about wind, rainfall, etc. and send this information to the company's server or in some cases to a cloud to be retrieved later. Figure 2—3 illustrates such an environment.

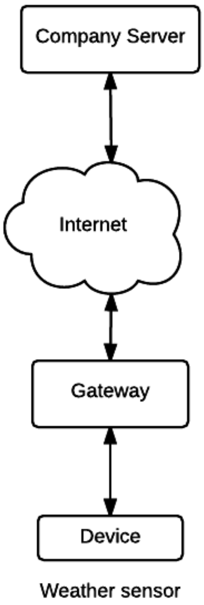


Figure 2—3: **Example of de-perimeterised environment**

To obtain Information Security in IoT it is required that systems and data are capable of protecting themselves *without* relying on basic network protection, such as firewalls. Firewalls effectively work as a perimeter to secure company resources from intruders, which in most cases are irrelevant for IoT. To simplify the deployment of more "things", these things must be able to enforce their own security policy levels (for applications, network access, devices, and individuals) even in an un-trusted environment or network. Another requirement is that the security mechanisms are simple, scalable, and easy to manage which simplifies the determination of their limitations since not all solutions fit in all environments [27].

The following techniques are required to embrace the de-perimeterised architecture:

- Security policy enforcement system
- Identity and rights management systems
- Encryption of data

2.2.1 Security threats associated with the Internet of Things

IoT security issues mainly consist of and are easily divided into two areas: virtual (see Table 2—4) and physical (see Table 2—5) threats. The physical threats increase as the things become more and more de-perimeterised. The virtual threats are closely coupled with the threats in any other IT-environment today and mainly consist of obtaining data and information (an asset) or taking control of the device itself. Additionally, applying the methods used for securing an IoT-environment are limited as many devices are constrained when it comes to performance and power.

Since this thesis mainly concerns the concept of Information Security, the starting point of the threat analysis has been the asset itself, which is information (data). Nor has a threat agent been identified since this analysis considers more general threats rather than specific ones.

By looking at the different points of attack it is easier to identify which threats are connected to IoT and also what vulnerabilities needs to be countered in order to secure each and every part in an IoT environment. The three identified points of attack are: the communication that occur between objects (IoT devices), the IoT devices themselves, and in the third case when a gateway is used, the central collection point of several sensors or a controller for several actuators.

Table 2—4: Potential virtual threats for any IoT environment

Virtual Threats affecting Information Security in IoT [36]			
Asset	Data & Information		
Point of Attack	Communication	IoT device(s)	Gateway
Threats	Interference (Denial of Service) Signal interception (Man in the middle) (Privacy Concerns)	Intrusion Exploitation (Privacy Concerns)	
Vulnerabilities	Uncontrolled or unprotected traffic flow	Insufficient authentication or authorisation Insecure user interfaces Insecure network services Insecure software/firmware Unprotected data	
Impact/Consequences	Compromised data Data loss Communication loss Inaccessible data Lose control of device	Compromised data Data loss Data corruption Inaccessible data Communication loss Lose control of device	
Information Security concepts affected	Availability Confidentiality Integrity Possession	Availability Confidentiality Integrity Possession Accountability Authenticity	
Countermeasures	Encryption of transport data Keep identification (IP-address) hidden	Reviewed applications, hardened operating systems, detailed traceability Secure environment and routines for development Security analysis and verification by third party The network uses strong encryption and signing Secure routines for physical access, log analysis, administration	

Table 2—5: Potential physical threats for any IoT environment

Physical Threats affecting Information Security in IoT			
Asset	Data & Information		
Point of Attack	Communication	IoT device(s)	Gateway
Threats	Interference (Electromagnetic compatibility)	Power loss Network loss Theft Sensor/device modification/ replacement	
Vulnerability	Wireless communication	Physical access to device Insufficient authentication	
Impact/Consequences	Communication loss Inaccessible data	Communication loss Inaccessible data Data loss Data corruption	
Information Security concepts affected	Availability	Availability Possession Integrity Confidentiality	
Countermeasures	Alternative (wired) network connection	Alternative power source Alternative network connection Move device to inaccessible area Authentication	

2.2.1.1 Communication

In communication there are two types of virtual threats that can arise, see Table 2—4. Interference occurs when the traffic flow (data) meant for the link, in some way is disrupted or totally eliminated because of other unwanted traffic flows occupying the physical link. A practical example of this is when a denial of service attack occurs – this is an attempt to make a machine or network resource unavailable, which can be devastating in an IoT environment that requires permanent communication [37]. Interference can also be done on a physical level, for example by jamming the wireless communication between nodes [38].

Signal interception can be done at several steps in the communication chain depending on what device or signals the attackers are able to listen to, for example sensors, actuators, gateways, etc. In the process of actually sending data a man in the middle attack could be done to secretly relay and in some cases alter communication between two parties [39].

Privacy concerns occur when attackers are able to interpret personal data because of insufficient authentication, unprotected traffic flows, or insecure network services both in communication and on the device level [40].

2.2.1.2 IoT-device

When looking at the IoT-device there are mainly threats in form of intrusion and exploitation at the virtual level. Intrusion is possible when an attacker utilizes security holes in insecure user interfaces, software/firmware or network services. There is also a possibility of intrusion when there is insufficient or non-existing authentication and authorisation for accessing a system, device, or data.

Exploitation occurs once a user has access to a component (device or gateway) in an IoT-environment. This exploitation can be in the form of reading further information, destroying data, or interfering with the communication by this component or others like it. For example, if the attacker is able to authenticate against the system it is likely that he/she will be able to get access to some sort of functionality of the device. This is why access control to limit the access rights of users is very important.

Physical threats can affect confidentiality, integrity, and availability in those cases when an attacker has physical access to an IoT-device or gateway. By tampering with or replacing a device it is possible, depending on the specific case, to do everything from reading or modifying to falsifying data. Therefore, limiting physical access and building in tamper resistance capabilities are a very important part of securing many IoT-environments.

2.2.1.3 Gateway

Other than those threats explained in Section 2.2.1.2 (IoT-device) a hijacked gateway can have a more extensive impact on a system since it is likely to have many sensors or devices communicating via it. The possibility to alter large amounts of data or making all or some of these devices/sensors unavailable are only few examples of the threats possible via a hijacked gateway. Note that these threats do not include all of the possible threats, but represent only a few potential threats that have been identified.

In order to maintain the dynamics of IoT it must be easy to deploy new devices, replace broken ones, and remove/ignore those that are unnecessary or have malfunctioned. That is why ensuring mutual authentication in end-to-end or hop-by-hop communication is extremely important. In this way unknown or false devices and gateways (deployed by an attacker) will be unable to communicate as they will not be able to authenticate against the current environment. This will limit some of the potential physical threats in terms by blocking an entry to the system and hence preventing further exploitation that could in some way affect information security.

End-to-end communication means that only the communication end points are able to see the data in clear text. This means that you do not need to rely on the immediate nodes to provide confidentiality or authenticity when sending data. For example, a gateway need not see the clear text of the messages that it is forwarding. The disadvantage is that the endpoint needs to manage all of the keys of the devices with which it is communicating.

With hop-by-hop protection the data sent from a device is decrypted when it comes to the gateway, and then is encrypted with the gateway’s keys before being forwarded. In this way only the gateway needs to manage the keys for those devices (for which it is responsible). The other endpoint only needs to manage the key for each of the gateways that it communicates with. The means that the gateway needs careful assessment in terms of code review, penetration tests, etc. in order to ensure that it will not expose the data or the keys.

2.2.2 Most relevant characteristics for securing Internet of Things

From the previous section it is clear that the most important things in IoT are mutual authentication and a way of securing the communication from each of the “things”. In order to utilize a collection of things there is a need for trust between the information sources and sinks. Trust is needed to rely on the information that each and every thing transfers. When communicating we need to ensure that the data is correct, which begins with being assured that we are in fact communicating with the “correct” device and that the data send by this device not has been altered on the way to its destination, i.e. to ensure integrity.

This is why we will focus on the first step in this process of actually setting up an IoT-environment: interconnecting the things. We want to be able to an assurance that you are talking to the correct device(s), hence you can start to build a chain of trust.

2.3 Identity and key management

A system needs to determine the identity of a thing with high assurance when this thing is seeking access or is going to be accessed (hence there is a need for mutual identification and authentication). As a result, there are two main problems concerning identity management in IoT:

- Identification – How do we uniquely identify *each* device in a secure way?
- Authentication - How do we authenticate each of the *identities* in a secure way?

The only way of knowing who or what you are communicating with is by identification[*]. Identifying a thing (sensor, device, gateway, or server) in IoT can become somewhat difficult in regards to security due to their restricted resources and communication methods. Also, the lifetime of an IoT identity and a given thing’s ownership are in many cases very short and there often exist identity relationships between several parties, i.e. users, administrators, or manufacturers that further complicate the process of identification and authentication. Pavel Legonkov and Vasily Prokopov examine the case of how to transfer ownership and control in the case of small cellular base stations in their Master’s thesis project [41]. Additionally, identifying a device in a secure way increases the possibilities of knowing whether it has been tampered with or not.

When a user is involved in an authentication process, a strong method of authentication is two-factor authentication (2FA). This authentication process exploits multiple factors which are based on the three following proofs:

- Something you have
- Something you know
- Something you are

The authentication that needs to be done between identities (things) cannot rely on a classic authentication mechanism such as usernames and password. Most things have to prove who they are by some sort of lightweight token or certificate (e.g. the private key of the certificate), which will be done by something that the thing *has* [42]. Minimizing the data sizes sent and (possibly) stored for this purpose is very beneficial for constrained devices.

The problem with “something you know” is that this knowledge (password or other information) needs to be programmed and stored, which often leads to people finding it. This also applies to something you *have* since it can be stolen as well, both logically and physically.

[*] Dealing with anonymous devices is outside the scope of this thesis and remains for future work.

Some of the basic identification methods of today are:

- URL (DNS) – associated with device IP – network level
- Software security tokens – Digital certificates (cryptographic keys), shared secret (configuration file, passwords)
- Hardware security tokens – internal or external modules generating one-time password (string, number, etc.)

Many of these are methods involve permanently storing some kind of data on the device in non-volatile memory (read-only memory, floating gate technologies, and fuse-based storage mechanisms) and are thereby exposed to both logical (virtual) and physical threats. Even though these physical attacks need extensive knowledge and tools, physical attack will with high probability be a bigger problem in the future, especially in correlation with massive deployment of IoT - when these devices are easily accessed physically. Even if the data is encrypted, simply limiting access to the cryptographic keys or secrets will not protect it. Therefore, it is essential that the keys used to protect the information that is used for authentication remain completely secret in order to guarantee a high level of belief in the authentication [43].

In order to address these types of attacks tamper-resistant microprocessors are a way of protecting sensitive information on a device, such as private keys. An example of such device is a Hardware Security Module (HSM). HSMs securely manage, process, and store cryptographic keys inside a hardened and tamper-resistant device with help from well-defined software-interfaces and physically protected hardware components [44, 45].

In some cases, it is possible to identify and authenticate an integrated circuit (IC) from its delay characteristics, even though multiple chips have identical digital logic functionality and are produced in the same way. In 2003 a report [46] written at MIT presented this possibility and proved it to be viable, but there were still some problems that needed to be addressed. This means that there are possibilities that things could prove who they *are* in some extension. The same conditions also enable implementations of physically unclonable functions (PUFs) which consist of a challenge-response mechanism to identify ICs uniquely which means that the function can be used produce an “unclonable” key for communication [47]. This is known as Hardware Intrinsic Security (HIS).

2.3.1 Identification with Hardware Intrinsic Security

As mentioned above, cryptographic security usually revolves around a secret key that needs to be stored somewhere in a system (device, gateway, server, etc.). HIS offers a way to address this problem, which requires that:

- A key should not be permanently stored.
- A key should be generated only when required. After use the key should be removed from all internal registers and memory. The circuit should be designed so that the former presence of the key does not leave a single trace when device is powered off.
- The key should be uniquely linked to each device; hence no reproduction or cloning should be possible.

If we can realize such a subsystem, then there is no need for expensive protection, e.g., using a HSM.

2.3.2 Physically Unclonable Function

PUF is a function embodied in a physical structure that consists of random characteristics originating from uncontrollable process variations during manufacturing, which can be described as a fingerprint of a device. These random characteristics could be sub-micron process variations (threshold voltage or gain factor) in static random access memory (SRAM) for example. Intrinsic-ID released a report [48] analysing different kinds of SRAMs and came to the conclusion that, even though test results vary, a suitable fuzzy extractor can be implemented to deal with the worst case situations by error correcting start-up pattern variations.

The following requirements need to be considered when producing a PUF [43]:

- **Low cost**: easy to implement with standard components.
- **Resistance to physical attack**: an attack meant to find out the behaviour of the structure should cause damage to it. The functional behaviour of the PUF should also change to such an extent that tampering is detectable.
- **Reliable**: produce low noise in the PUF data under different conditions, such as temperatures, humidity, and electromagnetic radiation.

Using a PUF usually consist of the two phases: Enrolment and Reconstruction. The enrolment phase (see Figure 2—4) occurs only once and is done when a new key or other information is being logically "stored". By generating an activation code from a key together with the device's PUF data, you can later use the same code and data to retrieve the key (reconstruction). The activation code is not secret and can be stored in non-volatile memory [49].

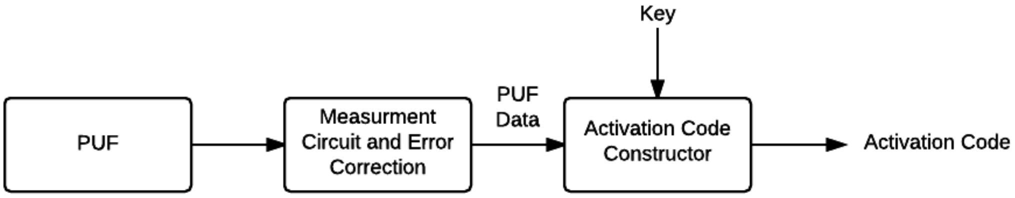

Key
PUF
Measurment Circuit and Error Correction
PUF Data
Activation Code Constructor
Activation Code

Figure 2—4: Enrolment phase of PUF

When the key needs to be used, reconstruction (see Figure 2—5) of the key is done. The generated activation code together with the PUF data is input to the Key Extractor which outputs the correct key – which should match the key from the enrolment phase [49].

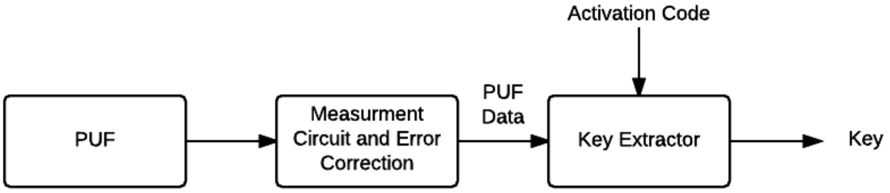


Figure 2—5: Reconstruction phase of PUF

Unclonability comes from the fact that you would have to know locations of these measured properties in the system with very high accuracy to be able to create a copy. Even if you know the properties, an exact copy will be very hard to reproduce, since even the exact same manufacturing process of an IC will result in differences between the devices. In those cases were an attacker tries to copy the PUF and use an activation key of another device, the key extractor will give an unusable key since the PUF data generated will not match the activation code [43].

Therefore, reconstructing keys from a "device fingerprint" gives us a unique identity that can subsequently be used when trying to authenticate a device. With this method of HIS we can ensure that our secret keys are physically secure and therefore we can have a greater trust in the identity of this device when communicating with a specific device.

2.3.3 Using PUF for identification

Now that we can use PUF to enrol a device and later reconstruct a key, we now consider how this can be used to provide secure authentication (keeping in mind the desire to avoid storing the key (as described in Section 2.3.1). To achieve a highly trusted identification of an IoT device we need to be able to send a random challenge and have it encrypted using the device's key (produced as shown in Figure 2—5) and then return the challenge encrypted with this key. The returned result has to match the result of encrypting the same challenge with the device's key.

As per the above we have shifted the requirement for secure storage of the key from the device (which now only needs to store the activation code – which is not sensitive data) to secure storage of the key at an authentication server which can perform the computation and match described in the previous paragraph. Note that this authentication server will need to be very secure, because if its copy of the keys for all of the devices for which it is responsible is compromised – then all of the devices can no longer be securely authenticated! Moreover, to produce new keys for each of the device will require that all of the devices have to be re-enrolled using a new activation code. This is potentially a very expensive operation as device would all need to be removed from their usual operating environment and placed into a secure environment for the re-enrolment. In the process we face an additional hurdle in that these devices cannot be securely identified! This occurs because potentially the only security we had for ensuring the device's identity has been compromised and anyone who leans the device's private key can create new device that would "impersonate" the actual device. This suggests that each device will also have to be externally marked in such a way that we can with high confidence identify it again using another means, which can for example be done with deoxyribonucleic acid (DNA). Of course this marking could possibly also be compromised – and the cycle of problems identifying the device continues.

2.4 Authentication and trust

A Public Key Infrastructure (PKI) is a concept that describes a set of technical mechanisms, procedures, and policies for the creation, management, storage, and revocation of digital certificates. A PKI usually exist of the elements shown in Table 2—6.

Table 2—6: Elements in a PKI [50]

Element	Description
Certificate Authority (CA)	Acts as root of trust in a PKI. Provide services that authenticate identities.
Registration Authority (RA)	Certified by CA to issue certificates for specific uses.
Certificate Database	Store issued and revoked certificates on the CA or RA.
Certificate Store	Saves issued certificates or pending or rejected certificate requests on the local device.
Key Archival Server	Saves encrypted private keys in the certificate database for recovery.

PKI uses public-key cryptography (also known as asymmetric-key cryptography) which consists of a public and a private key. The public key is used with other information such as subject, serial number, and version to create a data structure (called a “certificate”) [51]. This certificate is signed by using the CA’s private key. As a result the integrity of the certificate can be confirmed by decrypting this signature using the CA’s public key. This requires access to the public key for decryption and now the problem lies in how to authenticate the CAs public key. This is solved as the top of the hierarchy CA (root CA) uses self-signed certificates (i.e. certificates signed by the CA itself) since there are no other higher CA. That means that the whole trust is based on this certificate being legitimate.

2.4.1 Trust models

The number of organizations implementing PKI is increasing. As result there is a need for a trust model to interconnect different objects and identify them. Two primary trust models are used today to help address this problem: certificate hierarchies and cross certification [52].

In a certificate hierarchy the root CA delegates authority to subordinate authorities which can, in turn delegate authority to their subordinates [53].

Cross Certification enables trust between entities in different PKIs. This requires both CAs to issue certificates to each other. The path of trust (also known as a certificate chain – see the next subsection) that is created in this trust model is not hierarchal although the separate PKIs may be certificate hierarchies [54].

Section 4.2 purposes a method for authenticating identities and a trust model for use with IoT.

2.4.2 Certificate chains

A certificate chain is a hierarchal collection of certificates that leads from an end certificate back to a root of trust, i.e. the root CA. By verifying each of the certificates in a certificate chain, a party can gain trust in an end-entity. This requires that each certificate in the chain is signed by the public key of the certificate above it in the hierarchy (or the CA that is trusted in the case of cross certification), that this certificate has not expired, the certificate has not been revoked, and that the certificate conforms to the policies sets by the prior certificates [55].

Note that certificates can often be cached – hence the communication needed at run-time can be reduced. There are also techniques to do certificate chain reduction that can be used to reduce the effort required to process certificate chains (see for example [56]).

2.5 Internet Protocol Security

Internet Protocol Security (IPsec) is a framework for ensuring private and secure communication over IP networks. The protocol provides a number of functions and is quite flexible since it can offer functions such as:

- Access control,
- Connectionless integrity,
- Data origin authentication,
- Protection against replay attack (optional),
- Confidentiality (via encryption), and

- Protection for all protocols that may be carried over IP, including IP itself (through encapsulation).

IPsec uses two methods (protocols) called Authentication Header (AH) and Encapsulating Security Payload (ESP). Both protocols offer integrity (optional for ESP), data origin authentication, and anti-replay protection (optional), and with ESP confidentiality is also added. Both offer access control through cryptographic keys [57]. Setting up an IPsec connection involves many kinds of cryptographic methods, although most connections use two (rarely three) at a time. Authentication provides integrity through the use of an Integrity Check Value - a hash that is computed over a IP packet's content. The Integrity Check Value is computed through cryptographic hash functions such as MD5, SHA-1, SHA-2, or SHA-3[*]. The use of a secret key which is used by both parties allows the receiver to compute their own Integrity Check Value and hopefully get the same value.

ESP's encryption uses a secret key to encrypt the data before transmission which hides the actual content of the packet, thus preventing eavesdroppers. The encryption is done with a symmetric-key algorithm, such as DES[†] or AES. Symmetric-key algorithms use the same key for encryption and decryption.

Keys are provides in three different ways: manual key distribution (pre-shared keys), Internet Key Exchange (IKE) [58], Kerberized Internet Negotiation of Keys (KINK) [59].

IPsec can be run in two different modes: tunnel mode and transport mode. Tunnel mode acts like a standard tunnel and encapsulates the original IP packet inside a new IP packet thus creating a virtual tunnel between two points. Tunnel mode provides secure communication across the Internet or any other untrusted IP network. Transport mode protects end-to-end communication through authentication or encryption (or both).

The authentication header does not offer any encryption, but instead authenticates IP traffic. It authenticates the sender, receiver, and data, which protects against alterations and replay attacks.

Both the AH header and ESP header have some fields in common, specifically: Security Parameter Index (SPI) and Sequence Number. The SPI is used by the receiver to identify which security association the incoming packet belongs to. Each SPI can make a separate choice of cryptographic algorithm and of course has its own cryptographic data. The sequence number is used to assist against replay attacks.

Using AH in transport mode makes some modifications to the IP header and the AH header is inserted between the IP header and its payload (we have used a TCP segment as the payload in the examples), see Figure 2—6. Upon arrival the packet is authenticated and the AH header is removed and the IP header returns to its original state.

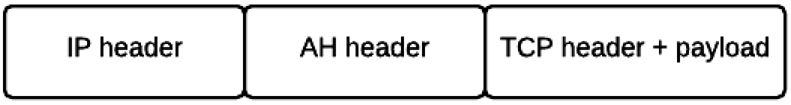


Figure 2—6: Example of AH used in transport mode

[*] U. S. National Institute of Standards and Technology (NIST). Microsoft, Google, and others have announced plans to deprecate MD5 and SHA-1 (with varying dates – but in most cases by 2017).

[†] DES has been depreciated by NIST since 2005, see RFC 6649 [97].

ESP in transport mode is a bit different since it adds some additional headers after the payload as well, see Figure 2—7. The authentication part is optional, i.e., it may or may not be added.

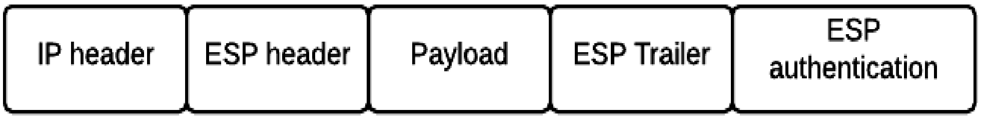


Figure 2—7: Example of ESP used in transport mode

To better explain what is authenticated and encrypted Figure 2—8 shows an example to better explain how these protocols differ. The grey colour in left hand figure in Figure 2—8 shows which parts of the packet are authenticated and integrity protected by the Authentication Data. In the IP header all fields except those that are mutable are integrity protected. Mutable fields are fields which may change value during transport, such as the Time-To-Live (IPv4) field and the Hop limit field (IPv6). By avoiding these mutable fields, we avoid the expected modifications to these fields leading to a different Integrity Check Value when the package would be authenticated. The next field in the IP header indicate which type of header comes next. When transforming the packet back into its original state the value in the next field inside the AH header and the ESP trailer replaces the protocol value in the IP header [57].

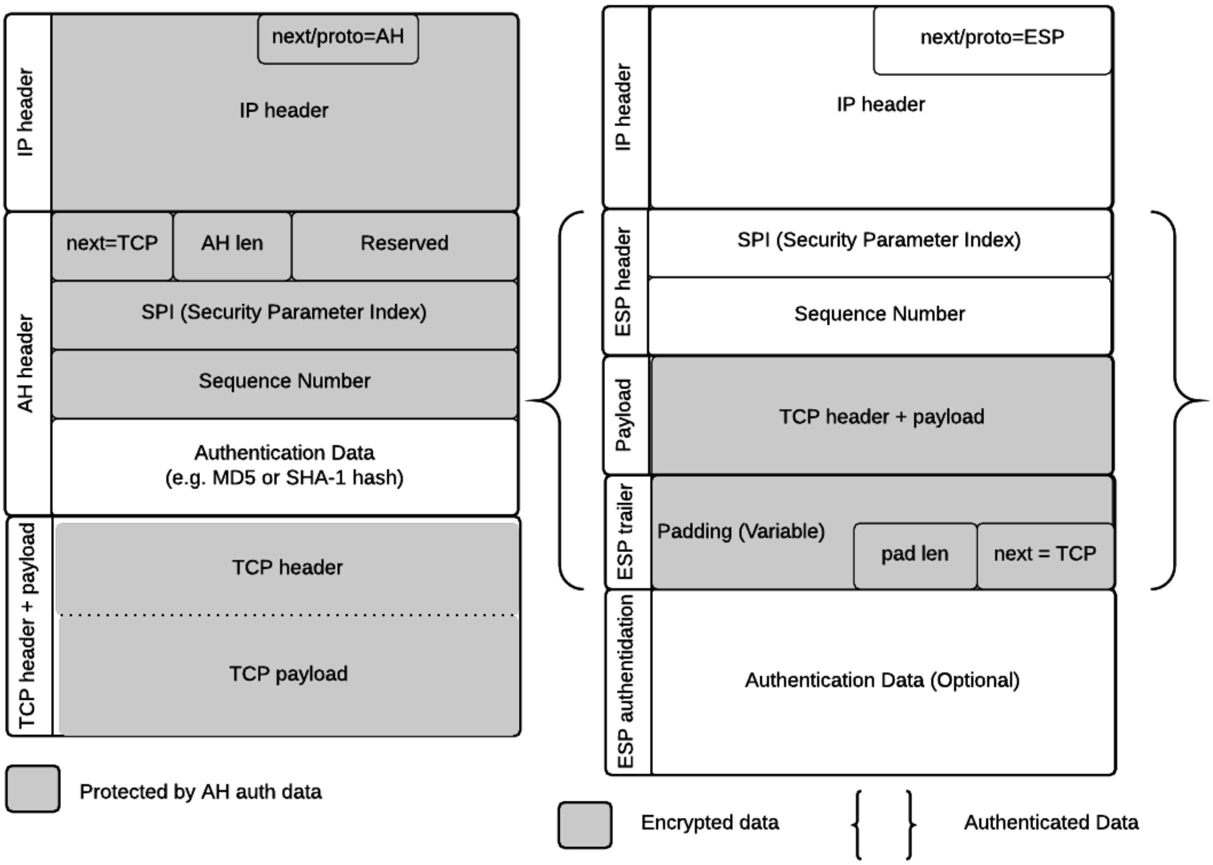


Figure 2—8: Detailed example of AH and ESP in transport mode

The difference between AH and ESP is that ESP encrypts the TCP header and payload *as well* adds an ESP trailer which describes the payload. As stated earlier, authentication is optional for ESP, but if it is used, then it is added as a final step. This means that the payload is first encrypted and then a hash is computed of the encrypted payload, ESP header, and trailer. Authentication in ESP does ***not*** authenticate the IP header, as illustrated in the right hand figure of Figure 2—8.

The essential differences are that AH provides integrity of the entire IP packet, while ESP ensures that the actual payload is confidential and optionally integrity protected.

In tunnel mode the structure of the headers is similar. The difference is that a new IP header is added and the original IP header and payload are appended following the AH/ESP header and then authenticated or encrypted depending on which protocol is used [57]. Figure 2—9 shows the structure of headers in AH and ESP mode. The actual encryption and authentication is identical in tunnel mode and transport mode, the difference is the addition of a new IP header in tunnel mode.

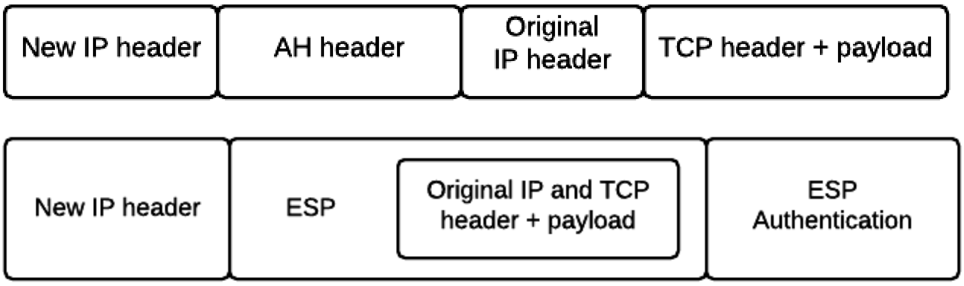


Figure 2—9: Example of AH and ESP in tunnel mode

As mention before the ESP and AH headers have a Security Parameter Index that indicates the relevant security association (SA) which the incoming/outgoing packet belongs to. A SA is used to specify which security properties are used by the host for this communication flow. This is the method which IPsec uses to keep track of all the information about a given IPsec communication session.

A SA may include attributes such as:

- SPI,
- Cryptographic algorithms for hashing and encryptions,
- Traffic mode,
- Encryption key
- Other parameter for the network (e.g. size for the Maximum Transfer Unit).

A SA is a one-way relationship between two or more hosts. For two-way communication each host needs two SAs, one for incoming traffic and one for outgoing traffic. In a system using IKE or KINK a Security Association Database (SAB) is used to store all the SAs. IPsec also uses a Security Policy Database (SPD) which defines policies for outgoing and incoming traffic. Policies define which types of packets should be dropped, forwarded, or accepted with/without IPsec protection. Policy decisions are made by checking different selectors which is very similar to a packet filter. These selectors can include (source or destination) IP addresses, user or system ID, transport layer protocol, SA, and destination and source ports. These databases are consulted to decide how to handle inbound and outbound traffic. For outbound traffic the SPD provides a pointer to the associated SA in the SAD, see Figure 2—10. For incoming traffic the SA is consulted in order to process the packet (including decrypting it). The packets are matched against selectors for incoming traffic to verify that the same SA is defined in the SPD [57] - if the SA does not match, then the packet is dropped.

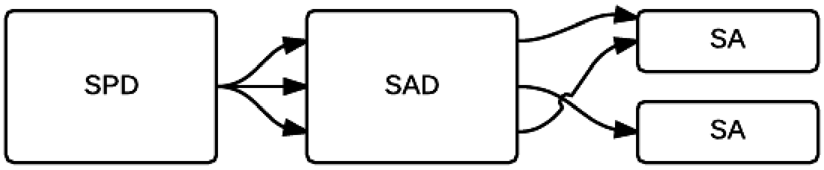


Figure 2—10: **Example of relation between databases**

When manual key deployment is used, both the pre-shared key and SA attributes are pre-configured by a user. This method is not scalable, but works well for small and static environment [57].

2.6 Transport Layer Security

The primary goal of Transport Layer Security (TLS) is to provide confidentially and data integrity between two communicating applications. TLS is implemented between the application layer and the transport layer. Encryption and decryption in TLS is similar to IPsec since it is based on symmetric-key algorithms. Note that as in IPsec, encryption is not mandatory. TLS consists of two main protocols: the handshake protocol and the record protocol [60].

The record protocol establishes rules for how to break up the data that is to be transferred. The record protocol is responsible for compressing data, applying a message authentication code (MAC), encryption, and transmission. The record protocol is also responsible for processing arriving data which is decrypted, verified, decompressed, and delivered to the application. The handshake protocol and two other protocols use the record protocol [60, 61, 62].

The handshake protocol negotiates a session between two TLS peers. This negotiation involves making decisions about many issues - including which encryption algorithm and MAC algorithm should be used. TLS has two other protocols which are used in connection with the handshake protocol: the Change Cipher Spec protocol and the Alert protocol.

The Change Cipher Spec message is sent by both peers in order to notify each other that following messages will be protected using the negotiated algorithms and parameters. The Alert Protocol is mostly responsible for conveying messages to another peer explaining the cause of a protocol failure [60].

A normal TLS handshake is shown in Figure 2—11. The Client and Server exchange a “Hello” with each other, thus establishing which protocol version is to be used, session ID, cipher suite, compression method, and two random numbers (one from sender and one from receiver). The cipher suit indicates which authentication algorithm, key exchange method, encryption cipher, and hashing algorithm are preferred by the client. The server’s reply contains the choice of cipher and a random number. Following these the server will sends its certificate, an optional CertificateRequest, (for mutual authentication), a ServerKeyExchange which is used to convey cryptographic information in order to complete a key exchange, and a ServerHelloDone indicating that the ServerHello is done. The Client sends corresponding messages together with a CertificateVerify to provide verification of its certificate. Then it sends its last two messages, ChangeCipherSpec and Finished, to verify that the key exchange and authentication were successful and that all of the following messages will be protected.

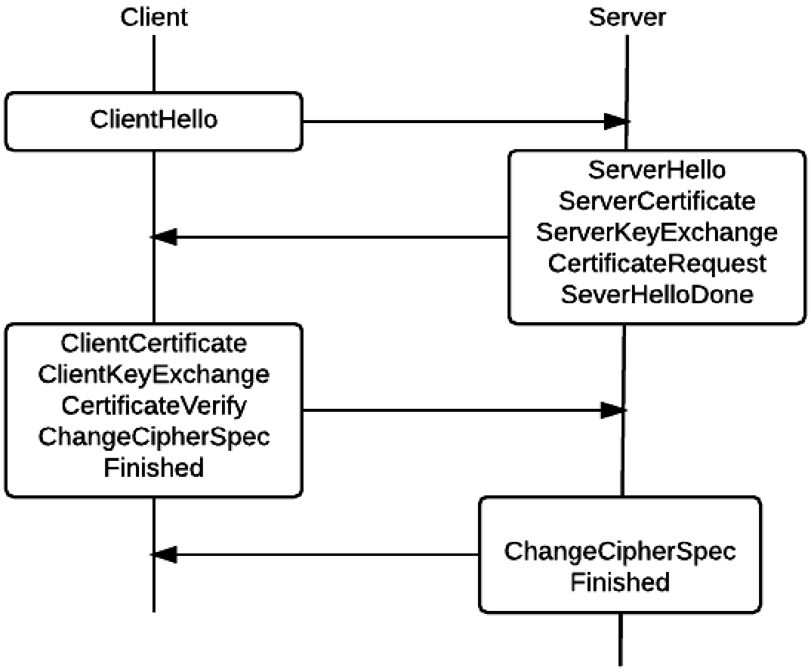


Figure 2—11: TLS handshake

TLS runs above Transmission Control Protocol (TCP) which is a connection-oriented and reliable protocol. TCP will ensure that packets are delivered and will resend any lost packets, unless the sender specifies otherwise. *"An acknowledgment by TCP does not guarantee that the data has been delivered to the end user, but only that the receiving TCP has taken the responsibility to do so."* - quote RFC793. TCP will also rearrange incoming packets in the correct order needed to deliver packets to TLS. Data sent with TCP is seen as a stream of data meaning that TCP does not with distinguish where one packet ends and another begins [63].

2.7 Datagram Transport Layer Security

Datagram Transport Layer Security (DTLS) is based on TLS. However, some modifications have been made to support datagram protocols (e.g. User Datagram Protocol (UDP)). Datagram protocols do not ensure that datagrams are received in order or received at all. These protocols are frequently used by applications that are either delay-sensitive or not sensitive to data loss [64].

Unfortunately, TLS is not designed to run over an unreliable communication channel, for example packets that arrive out of order cannot be decrypted since the integrity check depends on the sequence numbers. Thus if record number 2 arrives before record number 1, it will not be decrypted. DTLS introduces some reliability in order to ensure that the handshake process is successful. To do this it implements a retransmission timer in case of packet loss. Moreover, a queue is added to buffer handshake messages that arrive out of order. If needed, DTLS reorders them [64].

The major difference is that DTLS uses explicit sequence numbers for each record to enable DTLS to verify each individual record regardless of whether packets have been lost, arrived in the wrong order, or is a replay of an old message.

Unlike TLS, DTLS can be configured to not terminate a connection if a MAC fails; instead it simply discards the record [64].

2.8 IPv6 over Low-Power Wireless Personal Area Networks

Low-Power Wireless Personal Area Networks are networks which include devices and networks that fit the description made by the 802.15.4 IEEE Standard. These are often constrained devices, such as explained in Section 1.2, hence these devices may have limited CPU performance, network performance, memory, and power [65]. IPv6 over Low-Power Wireless Personal Area Networks (6LoWPAN) aims to extend and simplify the use of IPv6 through header compression and encapsulation so these IPv6 packets can be sent over IEEE 802.15.4 networks. In order to interconnect these networks there is a need for an edge router (gateway) that handles the compression/decompressing of IPv6 [66, 67]. AES is used to secure IEEE 802.15.4 networks on the link-layer. This works well for per hop-basis security, but is not well suited for end-to-end security [65, 66, 68].

Although 6LoWPAN was mainly developed for IPv6 networks, there are solutions for running this over an IPv4 network by creating a tunnel between the gateway and the host [67].

2.8.1 IPsec in 6LoWPAN

An extension to IPsec has also been developed in order to enable secure end-to-end communication between a sensor and a host [69]. This removes the need for a trustworthy gateway. These extensions compress the authentication header and the encapsulated security payload. By enabling these methods one can ensure the integrity and confidentiality of data that is sent over the sensor network and the Internet. The current draft only supports pre-shared keys for initial authentication, which means that SAs are manually configured. This is sufficient, but not very scalable since each connection will *initially* use the same key. Tunnel mode is not used when IPsec is implemented in 6LoWPAN since the purpose of compression is to make the packets smaller and using tunnel mode would only increase the packet size [69, 70].

2.8.2 DTLS in 6LoWPAN

In an internet draft a similar extension has been made for DTLS to enable its use in 6LowPan networks [71]. The purposed extension reduces DTLS Hello messages but also discuss the possibilities to omit some messages. This draft is also made using a cipher suite that is based on pre-shared keys [72].

Above DLTS a web protocol called Constrained Application Protocol (CoAP) was implemented. CoAP is a web protocol specialized for constrained nodes and networks which run over datagram transport protocols such as UDP [72, 73, 74]. Using CoAP together with IPsec has also been discussed in a similar informational draft [75].

2.8.3 6LoWPAN and interconnection with IPv6 networks

Unless the 6LoWPAN is a separate IPv6 subnet there is a problem in directly incorporating 6LoWPAN devices into an IPv6 network due to issues regarding IPv6’s neighbour discovery protocol. A solution to this problem is proposed in a Master’s thesis by Luis Maqueda Ara [76].

2.8.4 Privacy of IPv6 addressed nodes

It is important to remember that IPv6 devices can create their own IPv6 link local address based upon the IEEE 48 or 64 bit median access and control layer identifier that was assigned by the manufacturer. This process is called auto-configuration. However, this introduces a privacy problem, since each such node will always have the same lower order 64 bits in its IPv6 address. Thus it is trivial to notice that a given device is part of a network for each network that it becomes part of.

To void this problem the SEcure Neighbor Discovery (SEND) protocol has been introduced in RFC 3971 [77]. Unfortunately, there has been limited work to integrate SEND, 6LoWPAN, and IoT.

2.9 Summary

Information security is a term used to describe, among other things, processes for securing data and systems. An example of such process is a threat and vulnerability analysis, which can be used to identify threats and thereby identify and evaluate security solutions for an IoT-environment. From our analysis, a group of potential virtual and physical threats have been presented to give an overall picture of the needed security requirements. Following this, background information was given on secure key storage and authentication methods that have been proposed together with a summary of some of the most popular means of communicating securely.

3 Methodology

To be able to answer our questions (as proposed in Chapter 1) it was important that we had the knowledge required to answer these questions. This meant that feasibility studies are an important part of the project, because this would provide us with relevant knowledge to supplement what we already knew.

In order to answer our questions about security for IoT and its relation with constrained devices we had to review research which had been done earlier in order to avoid performing the same research twice. This is a good method in order to merge different technologies and to conduct out research, especially if we want to find a *practical* solution.

3.1 Research Process

Figure 3—1 shows some of the major milestones in this thesis project. During the first week we defined a more detailed method and planned how we would work during the following weeks.

Weeks 2-5 were to be used for feasibility studies through which we would deepen our knowledge within the subject and to see which information/technology already exists and then use that information to see what possibilities there are to further analyse and hopefully find places for improvements.

In week 5 we reached an important stage in our project. Here we evaluated our feasibility studies and decided how to continue the project. For example, we could have chosen a theoretical path and further study the subject or examine our alternatives and eventually provide a practical solution or do an interview based study. Further issues about an interview based study were discussed during the first week of practice.

If a practical solution is to be realized, then there is probably a need for some kind of hardware and/or software. The purpose of a practical solution would be to solve a problem that has yet to be thought about. By this stage we planned on having a first draft of our thesis ready.

We chose to do an interview based study for the following reasons:

- To gain insight about todays' business needs and views regarding IoT
- To gain insight in how one secures an IoT-environment
- To gain insight about general solutions for IoT-environments

The plans for these interviews will be described in the next section and the details of the interviews are given in Section 4.1.

A second draft is planned to be ready 1-2 weeks before the actual examination. After the final examination we will review what may need to be improved based upon feedback from our opponent(s) and questions at our oral exam. Within a week's time after the oral exam we will submit a final version of the thesis.

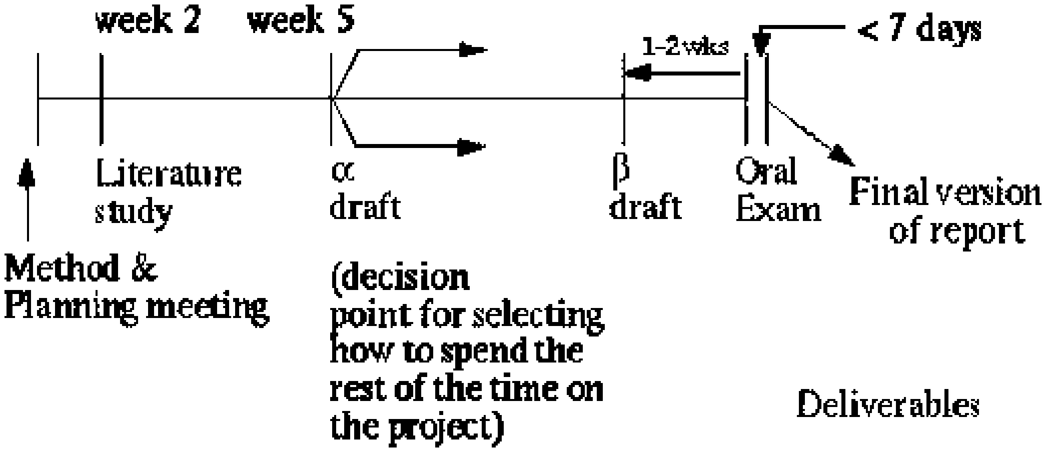

week 2
week 5
1-2wks
< 7 days
Literature study
α draft
β draft
Oral Exam
Final version of report
Method & Planning meeting
(decision point for selecting how to spend the rest of the time on the project)
Deliverables

Figure 3—1: Generic timeline (Used with permission of the author – our examiner.)

3.2 Interview study

The interview study is split into three parts. The first interview involves questions about how our clients (i.e. the expected clients of Combitech AB) view IoT with regards to area of use and expected threats. Based on these answers an IoT-environment was derived. This environment was used as input to a second interview where this environment was analysed in order to create a security design. For the third interview, the IoT-environment and security design were analysed in order to obtaining technical solutions for this particular environment and design.

All interviews were recorded, transcribed, and then analysed in order to summarise their content.

3.2.1 Interview questions

Here we map the different questions used in each interview. Note that the questions used in interview number 2 are based on the results of interview number 1. Interview number 3 has the same relation to interview number 2, because the questions asked during interview number 3 are based on the answers from the previous two interviews.

3.2.2 Interview number 1

The goal of the first interview was to gain insight about how businesses view IoT today. This was done by identifying what Combitech AB’s clients are in need of and how clients see on the IoT in regard to security. Note that the person interviewed was a sales person representing potential clients.

3.2.2.1 Needs

The questions used in order to answer questions about clients' needs were:

- What is the client's view of IoT and how do they define it?

  This question was used in order to understand how the client sees IoT and what it means to them. In relation to this, it is also interesting to see how a salesman defines IoT.

- Is there a specific area of use?

  The focus of this question was to see in which area(s) IoT is to be implemented. Is it the industry sector, private sector, or public sector? Another sub-goal for this question was to learn how the devices are expected to be used. For example, is the usage remotely controlled devices or is the intention to create half/fully automated solutions.

- What is the purpose of adopting IoT?

  The purpose of this question was to see what different client's goals were with regard to their adopting IoT. For example, is the goal to make a process more efficient or to develop new products/services?

3.2.2.2 View of IoT

Here the questions were focused on the clients' view on security of IoT and what concerns they have in regards to IoT security.

- What is the clients' view of security in IoT?

  The purpose of this question was to find out if the clients felt safe when using IoT and if they thought IoT was sufficient secure.

- Are there general concerns because they do not know the area or are they very optimistic/hopeful?

  This question targets the reason behind clients fear regarding IoT or if they do not fear the concept of IoT at all.

- Are there any general threats? Are there any specific threats?

  With this question we hope to learn if clients fear any specific threats. For example, do they fear that other people will have access to the information they send or do they fear that people will be able to exploit and sabotage their systems.

3.2.2.3 Interviewee

The interviewee responding in the first interview works as sales manager in the business segment: industrial products and services. Their role as a sales manager was fitting regarding questions about client needs and their views on security. This person is often in contact with different kinds of clients and these contacts provide him/her with useful knowledge and insight about IoT with regard to different clients' needs. How this may have affected the reliability and validity of information is explained in section 0 and 3.5.

3.2.3 Interview number 2

The purpose with this interview was to establish a design for an IoT-environment. The preparations before this interview involved developing an IoT-environment and an associated description. The description expressed demands about how the system was supposed to work. These demands were based upon the first interview. The IoT-environment and the description were sent to two interviewees so they could prepare themselves for the interview.

3.2.3.1 Questions

The actual interviews did not involve any direct questions, but instead took place as a free flowing discussion. This method was adopted because the purpose of this interview was to allow the interviewees to analyse and modify the design of the IoT environment. Based upon the description of the IoT-environment the interview answered sub-questions such as: how can the presented IoT environment be secured, which security is required in specific parts of the system, and how can the set of requirements be realised from a security perspective.

3.2.3.2 Interviewees

The interviewees in the second interview were two information security consultants working primarily with information security requirements, design, and reviews. An information security consultant often creates a so called "business analysis" which includes identifying business specific requirements, legal requirements, information requirements, and after performing a threat and risk analysis for the business the consultant presents a set of information security requirements. An information security consultant is also often involved in IT security architecture. This field of work provided valuable input for designing and analysing an IoT environment because it involves contact with clients and frequent work with business analysis and IT security architectures.

3.2.4 Interview number 3

The purpose of this interview was to identify the technical security solutions that could be implemented in the IoT-environment in order to fulfil the requirements which were derived from Interview number 2. Preparations for this interview involved establishing a security model for the IoT-environment based upon the information gained from the two people who were interviewed in the previous interview. The model included customer demands and security requirements. The goal was to identify ideas for technical security solutions regarding communication and storage in the proposed IoT environment.

3.2.4.1 Questions

Similar to interview nr 2 the interview took place as a free flowing conversation. Based on the environment derived from interview nr 1 and the requirements specified during interview nr 2, the interviewee analysed the results in hope of being able to provide suitable "technical" solutions. Technical solutions refer to the use of specific protocols, encryption, etc.

3.2.4.2 Interviewee

The interviewee works as an IT security consultant who examines and reviews system security designs to ensure that the purposed security solutions are implementable in the design based on hardware and software specifications. This involves evaluating solutions on a technical level in order to ensure that the desired level of security is fulfilled.

3.3 Literature study

The literature study was done with the questions in mind when searching databases for similar work. Information could be analysed and compared from various sources. The main search tools that were used were Ex Libris Ltd.’s Primo® and Google. Primo is a search tool provided by KTH which provides a single access point to everything available in the KTHB Library Catalogue. Primo offers various types of information such as books, journals, reports, conference proceedings, and research publications (from KTH and other universities) [78].

The second search tool that was used is Google [79]. Google was mostly used in order to identify both scientific and non-scientific information. The search terms used in this process were based on the questions written in Section 1.6 and questions derived from discussions between ourselves and our supervisors at Combitech AB. Starting from those questions, the results were analysed and evaluated. While evaluating the results further searches were made in order to answer the questions posed in Section 1.6.

Figure 3—2 shows a simple search process. The first step was to identify the purpose and goals for this thesis project. Then we identified and located potentially relevant sources of information which were then evaluated. Steps three, four, and five involved acquiring information, creating results, and presenting results to the involved parties in order to evaluate if the results satisfied the information needs. This is a cyclical process that iteratively acquired information which helped us complete a pending task.

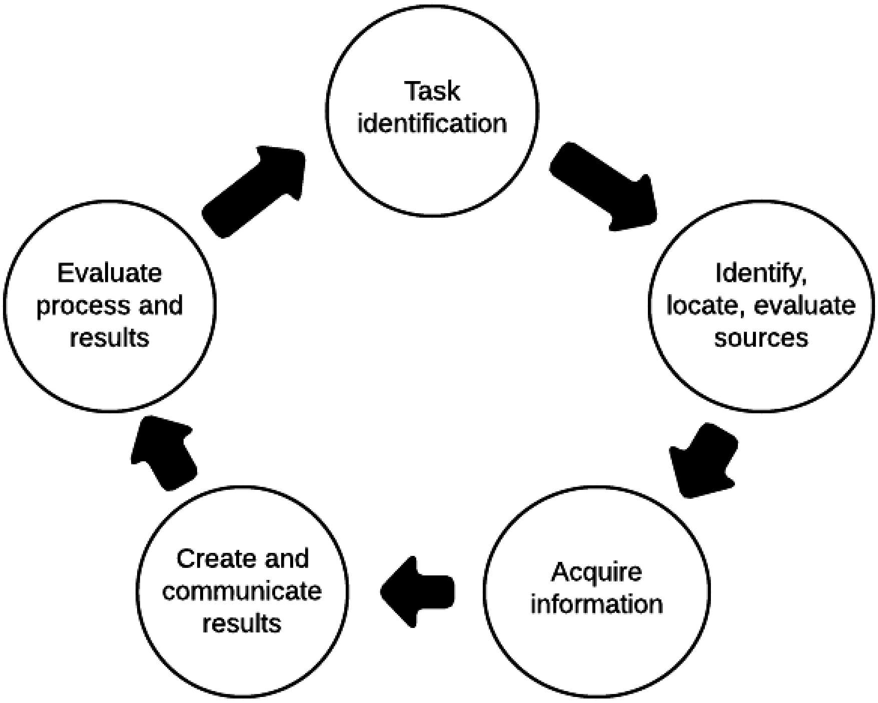

Task identification
Identify, locate, evaluate sources
Acquire information
Create and communicate results
Evaluate process and results

Figure 3—2: Simple search process

3.4 Reliability

This section explains how reliable the gathered information was. The information's reliability was analysed for two sources of information; interview study and literature study.

3.4.1 Interview study

Since the information gathered from the interviews are mostly based on the interviewees' business experience and personal knowledge there is a need for confirmation with other sources. In some cases, such as the first interview when asking about specific customer needs and thoughts about IoT security, this was not possible. The small number of interview participants means that the information provided is very dependent upon these persons' work experience and opinions. This means that the information's reliability is based on qualitative sources, rather than providing quantitative results. For example, during the first interview this person functioned as a relay for customers' opinions and needs, which means that we as researchers have to rely on this person to providing reliable information.

3.4.2 Literature study

The literature study was very similar to our interview study. Where multiple sources were available a comparison and analyse has been made to strengthen our source criticism in order to utilized to acquire reliable information. The literature study is mostly based on information provided by researchers, companies, and associations. When reading works produced by other researchers it is critical to compare similar kinds of research in order to ensure that their results are consistent with each other. This includes comparisons of opinions.

3.5 Validity

This section explains how valid the gathered information was. The information's validity was analysed for two sources of information; interview study and literature study.

3.5.1 Interview study

The information derived from the interviews is based on the knowledge and extensive experience from people who work within the field of information security. Therefore, we deem the information to be valid because the questions and IoT environment were formed in a way which relates to their normal field of work.

3.5.2 Literature study

The validity of the information found during the literature study was ensured by applying source criticism. We reviewed each source and ensured that the information provided was the recent. For example, when researching TLS it was important to read the latest Request for Comments (RFC) for TLS, thus ensuring that the information is up to date.

It was important to differentiate between the different types of information. Reading international standards does not involve as much source criticism as when gaining information from other sources, such as informational drafts or research papers. In the last two categories ensuring validity requires thinking more about the actual content.

4 A first step to secure the Internet of Things

Together with this chapter and the background information, our knowledge is applied to the three scenarios.

4.1 Interview

The goal of the first interview was to acquire knowledge about what Combitech AB's clients are in need of and how these clients see IoT in regard to security. The questions or structure used in each interview were presented in Section 3.2.1.

The goal of the second interview was to acquire a design for an IoT-environment. The preparations before this interview involved developing an IoT-environment and an associated description.

The goal of the third interview was to identify the technical security solutions that could be implemented in the proposed IoT-environment in order to fulfil the requirements which were derived from Interview number 2. The purpose was to acquire ideas for technical security solutions regarding communication and storage in the established IoT environment.

The IoT-environment and description which were presented to the interviewees during interviews 2 and 3 are presented in Appendix A and Appendix B (respectively).

4.1.1 Interview number 1

The interviewee stated that IoT is generally associated with connecting systems, machines, or products in some way in order to enable value for the business. IoT originates from the concept of machine to machine, which initially involved peer-to-peer communication and with people communicating with a machine (interface). These users' only possibility was to perform a query and read a response.

Today when business client discuss IoT it involves looking at existing standards and the possibility of including the systems, products, or machines in a larger context. This desire of operating in a larger context adds demands on all kinds of standardisation (interfaces, systems, etc.). There are many different types of clients in the sense of that some clients are less experienced than others. Some clients may be at a beginning level when discussing IoT and are simply talking about smart products, which primarily involve local solutions using an interface and locally presenting the information. This might be a first step before connecting a product and enabling the ability to view/set/manage a value via a smartphone, web interface, or app[*]. One of Combitech AB's focuses is effect (why one wants to do something and what they wish to achieve). Depending on which effect one wishes to achieve, a solution will require different types of actions where some clients need to make greater efforts than other in order to reach their goals. However, in order to reach the era of IoT the interviewee believes that there should be a plural somewhere, meaning that a system should be part of a larger system, thus creating a system of systems.

To summarise, the idea of IoT is seen as connecting systems, machines, or product in order to enable value and a final goal is to be a part of a larger context.

[*] Short for application.

When the interviewee was asked if there was a specific area where IoT was going to be used he/she preferred to talk about trades* instead. The trades where IoT could be implemented covered many areas, such as defence industry, medical technology, and classical production industry. Where IoT is going to be used depends on which trade the client is in.

Instead of seeing the area of use as a specific sector it can be seen as providing information and a basis for decision-making for a user, where the "user" can be a system or a human. Looking at the concept of systems of systems, the information that is provided by a system may be a puzzle piece in a larger context. The information itself can be split in two categories: operative and strategic information. Operational information is used as a basis for decision making, while strategic information can be collecting data for analysis. In regards to information the interviewee chose to use the term "Information of Things" instead of "Internet of Things" because even though Internet is involved in the process it is the *information* that is vital.

There are many reasons for implementing IoT. A potential goal might be to create an ecosystem around a product by "opening" it through Open Innovation. In this way a business can encourage others to develop services around the product. Another example is an automotive vehicle which requires collaboration between different sensors and connecting the vehicle to an ecosystem might involve collecting data from the local traffic administration in order to create autonomous routes. These autonomous routes can be used for an autonomous vehicle in a delivery service, i.e., where the vehicle drives itself and selects the currently ideal route.

The purpose of using IoT might also be the desire to improve or change an aftermarket or perhaps to make a service more efficient.

The interviewee explained an interesting example of General Electric, who stated that by making their fuel consumption 1% more efficient they would save USD$30 billion in 15 years. General Electric is working a lot with interconnection to allow for adjustments to more parameters than has been previously possible [80].

Security is often unresolved when it comes to the IoT. The majority of interviewees say that it is important, but only a few have answers of how one should approach the security challenges.

The security of IoT might be neglected if it is a simple product (e.g. a flat-iron), because one may not see the potential risks of something larger happening or that this device could potentially be used as entrance to something bigger (i.e. because the product is a subsystem of another system). The business clients that see the potential risks fear that their own system will become this entrance hence introducing a vulnerability to all other systems which most likely will be interconnected in the future.

A classical issue is that business clients do not desire security if it leads to limiting functionality. Hence it is important to find a balance between functionality and security. In the industrial market security is generally associated with problems because implementing security often involves additional efforts and costs. Depending upon the challenges one faces it is important to implement the appropriate security functions.

Another common scenario is that businesses often rely on a third-party cloud service because they are seen it as "safe". The problem is not the service provider, but rather that businesses miss the threats when information is traveling to/from the cloud service. That is why it is important to value the information rather than treating the information as non-critical. In order to value information one needs to explore what threats and risks there are to different kinds of information. This also relates to the importance of implementing appropriate security functions, but in order to

* Business or line of business is often a synonym for trade.

do this there is need to value the information. Depending on the kind of information it may be also be subject to laws (e.g. The Data Protection Directive regarding the processing of personal data).

During the process of developing a system it is important to identify threats and risks at an early stage. These threats and risks are used to identify requirements when implementing security in an IT-architecture. The car industry is an example where security is essential because a security flaw can lead to harm and in the worst case death.

As the systems become more complex so do the threats, the Stuxnet malware is a good example of such a threat [81]. Today in IT, threat agents are usually organised crime and foreign government, rather than a lone hacker which was more common earlier.

4.1.2 Interview number 2

It is difficult to find security functions that are specific for IoT. What has to be addressed is if there are possibly unique constrains upon the system and where should the security be implemented to avoid negatively affecting functionality. Analysing the sensitivity of data in a system is vital, since sensitive data might have to be handled differently in different phases (storing, use, and transport). It is also important to determine which phases are relevant to a specific system. For example, data that is collected from a sensor might never be stored locally; instead it might be transmitted as soon as it is collected. In this case there might not be any reason to consider the security of the information when stored locally – as the data will only be present for a short period of time.

When designing bigger systems trust becomes a big factor. There must be assurance that the correct system is providing the information. In the sense of trust, authentication becomes vital in order to identify devices and then building a trust model between them, not only for the individual devices but also for third-party devices. Regarding data sensitivity, it is important to place appropriate requirements on third-party devices, since they will most likely handle data in different phases.

Logging may also be relevant if a data intrusion occurs, since logs might provide the possibility of tracing the attack. However, IoT devices may be constrained and have limited storage so logging might require a central point to provide this storage thus enabling them to be analysed at a later stage. One should also be critical of what information to log, because logging all information is not always necessary. Additionally, in some setting some information may need to be logged for reasons of compliance, audits, and other regulations or for use with in litigation.

Systems in general place a lot of requirements on what software, hardware, and communication protocols are used in order to ensure that the system is secure. For example, it is preferable to use open and standardised protocols instead of proprietary protocols. However, in some areas no specifications or standardized protocols may be available. For open protocols there already is increased reliability by using something that is standardised, established, public, and proven (e.g. HTTP and HTTPS).

It is important to analyse the needed functionality of the devices based on use cases and restrict their functionality to a level where they can perform what is required, not more or less. In the case of a device that is just streaming data outwards, limiting the possibilities of sending traffic inwards might be a good way to increase the security. However, incoming traffic in the form of a software update may require authentication in order to confirm that the update is arriving from the right source. A master's thesis written by Mussie Tesfaye proposes a solution as well as discusses and evaluates existing solutions regarding “Secure reprogramming of network connected devices” [82].

Incoming traffic from a device which contains data might have to be compared with reference data in order to confirm that the data being provided is reasonable, thus creating a form of sanity check to confirm if data is likely to be correct.

Designing a secure environment involves many factors, such as developing trust models, authentication, securing communication flows, securing interfaces, etc. The analysis also involves physical security aspects. If the devices are physically exposed, this might put additional requirements on the software. Crisis management might also be a part of the design and analysis process, for example if a device goes offline, should it be assumed that something bad has happened to the device? Was the device stolen or is there disturbance in the communication?

Strong security is often complicated and therefore expensive, thus it is important to find an appropriate level of security in terms of safety, functionality, risk, and costs.

4.1.3 Interview number 3

To produce a practical solution a security audit needs to be performed based on a security goal or a threat and risk analysis of the system or environment. There is a need to know what kind of exposures that exists in a system to imagine what a potential attacker could exploit. You also need to know what kind of underlying hardware and software exists to determine if a solution is applicable to start with, that is, you need to know what things you are interconnecting.

It is required that the systems being deployed, fulfils fundamental IT security principles, such as:

- End points only run reviewed applications, hardened operating systems, and have detailed traceability;
- The network uses strong encryption and signing (e.g. TLS 1.2) and key handling using a PKI (with your own CA);
- The system is verified and examined by an independent party;
- Secure routines (e.g. physical access, log analysis, and account administration); and
- Secure development environment (e.g. anti-virus protection).

A system is never going to be completely secure. It is only a question of how hard and time consuming you want to make it for an attacker to exploit your system.

4.2 Identification and authentication (Trust model)

With a suggestion for a model to uniquely identify our devices, we will now propose a concept to authenticate them in order to enable communication and trust. The following subsections describe a proposed simplified way of managing identities and trust in an IoT environment using a Public key infrastructure (PKI). Note that this is a proposal and *not* an applicable solution. The proposal should be described in more detail and during this aligned with current best practices before implemented. Furthermore IoT devices that cannot store a certificate and the certificate of a CA are not considered in this section and are left for future work.

4.2.1 Trust

When a device or gateway is enrolled, the device's own certificate needs to be stored somewhere in the top level hierarchy server (CA). With the private key of the certificate safely reconstructed each time it is used with a PUF, it gets hard for an attacker to falsify this device's identity (i.e. retrieving the private key). A device or gateway also needs to know who it should connect to, that is, the gateway or server certificate needs to be known before deployment. This certificate can be pre-stored on the device together with certificates from the CA (where these certificates are for the device itself and for the CA).

Companies will always want to create their own services and sell them to other companies or people. That is why a chain of trust with cross certification (as explained in Section 2.4.1) is a very good way of delegating rights and authorisation to others by pushing information to already trusted devices.

The owner (company or person) of the environment will work as an authority for the whole trust chain created between gateways and devices. The Server/Cloud will create a trust between itself and the gateways. The gateways will in turn do the same with these devices connected to it. There are also possibilities for devices to connect directly to the Server/Cloud just as a gateway.

In the following sections the basics of identity management and authentication of nodes in an IoT-environment is shown. The names of the different certificates (GW/D/S/NS-IoTcert) are only used to clarify which identity is being sent. The certificate should contain at least a public key which is sufficient to identify a device.

4.2.2 Authentication of gateway

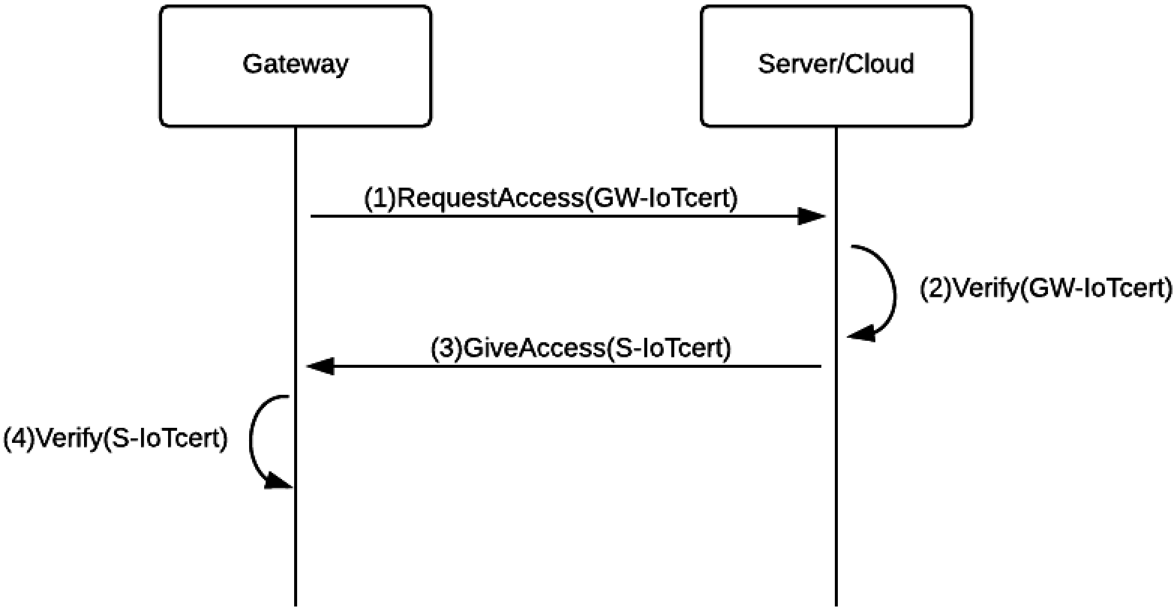


Figure 4—1: Authentication of Gateway

The identity (public key) of the Gateway needs to be stored in the Server and vice versa to enable proper authentication during the initial deployment. For example, when a gateway wants to authenticate a server/cloud it will carry out the following steps (shown in Figure 4—1):

(1) The gateway sends a message encrypted with the Server/Cloud public key and then signs it with its private key. This ensures that the gateway sent the message and only the server will be able to decrypt it.

(2) The server will be able to verify the gateways identity or not depending on which keys were used. The server tries verifying the signature with the public key of the gateway and then decrypting the message with its own private key.

(3) The Server/Cloud sends back a response only if it verifies the gateway’s identity. The response is a message encrypted with the gateways public key and thereafter signed with the Server/Cloud private key.

(4) The gateway verifies the server the same way as in Step 2 since the gateway knows the public key of the Server/Cloud as this was installed into the gateway pre-deployment.

When trust has been established between the server/cloud and the gateway, the server/cloud can then push device identities to the gateway for it to use when authenticating each of device.

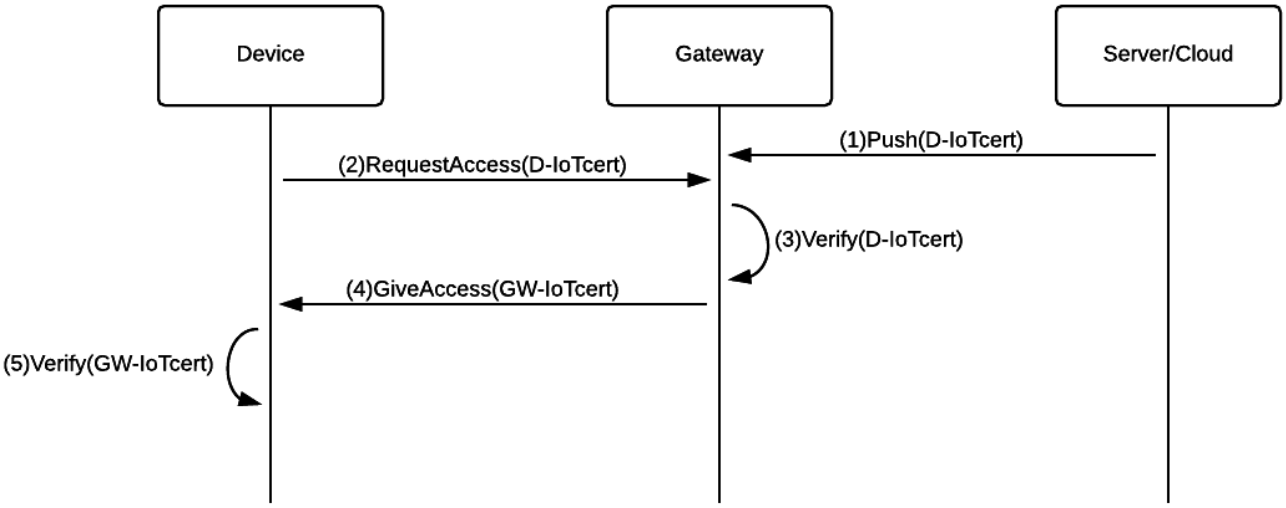


Figure 4—2: Authentication of IoT-device

4.2.3 Authentication of IoT-device

Since the gateway is now trusted, we can start authenticating devices against it. Each newly deployed device is assumed to have been pre-configured with the identity of its gateway; hence it has a copy of the gateway’s certificate. If it does not know who to authenticate with, then it could basically connect with anything. The process of authenticating IoT devices is shown in Figure 4—2 and consists of:

(1) The new device’s identity is pushed to the gateway from the server, so that the gateway can learn about the new device and hence can communicate securely with it.

(2) When a device connects to the gateway it sends a request for access to the gateway, same as Step 1 in Chapter 4.2.1.

(3) Since the gateway knows the identity of the device it tries to verify the identity of this device, same as Step 2 in Chapter 4.2.1.

(4) The gateway sends a response if it successfully verifies the identity of the device.

(5) The device verifies the gateway in the same way (as Step 3) since it knows the identity of the gateway prior to its deployment.

Since there now is a chain of trust in the environment we can manage, revoke, and give other systems access to our gateways or devices.

4.2.4 Revocation of devices and gateways

In those cases were we suspect that a device or gateway has been compromised it is easy to revoke its access by either disabling the device's access via the gateway (at the gateway) or by having the Server/Cloud pushes a revocation order for a specific device to the gateway (the latter is shown in Figure 4—3). In some cases the gateway may identify suspicious activity, such as a device trying to authenticate twice or that there are data deviations from sanity checks. In those cases the gateway could notify its supervising node in the hierarchy and then this node can decide what action is appropriate.

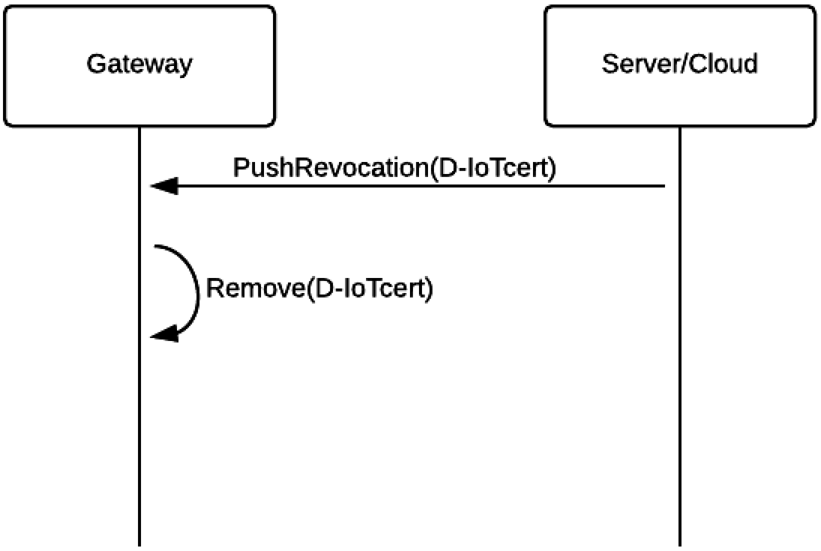


Figure 4—3: Access revocation of device

4.2.5 Access control

If another company wants access to our sensors or gateway for example, the following method could be used granting access. It could also be used to transfer ownership. We assume that the top of the hierarchy Server learns the identity (NS-IoTcert) of the New Server/Cloud (other PKI CA) requesting access. (This could be done in several different ways.) Now the process proceeds as follows (as shown in Figure 4—4):

(1) The Server/Cloud pushes the identity of requested gateway to the New Server/Cloud.

(2) The identity of the New Server/Cloud is pushed to the gateway that it wants to connect to.

(3) The gateway now waits for an access request from the New Server/Cloud.

(4) When a request for access has been received, the gateway verifies the sender with the identity pushed from the Server/Cloud.

(5) If the New Server/Cloud identities matched, the gateway sends a response to the new server.

(6) The New Server/Cloud authenticates the gateway with the identity given from Server/Cloud.

(7) A response is sent back to the gateway if the authentication was successful.

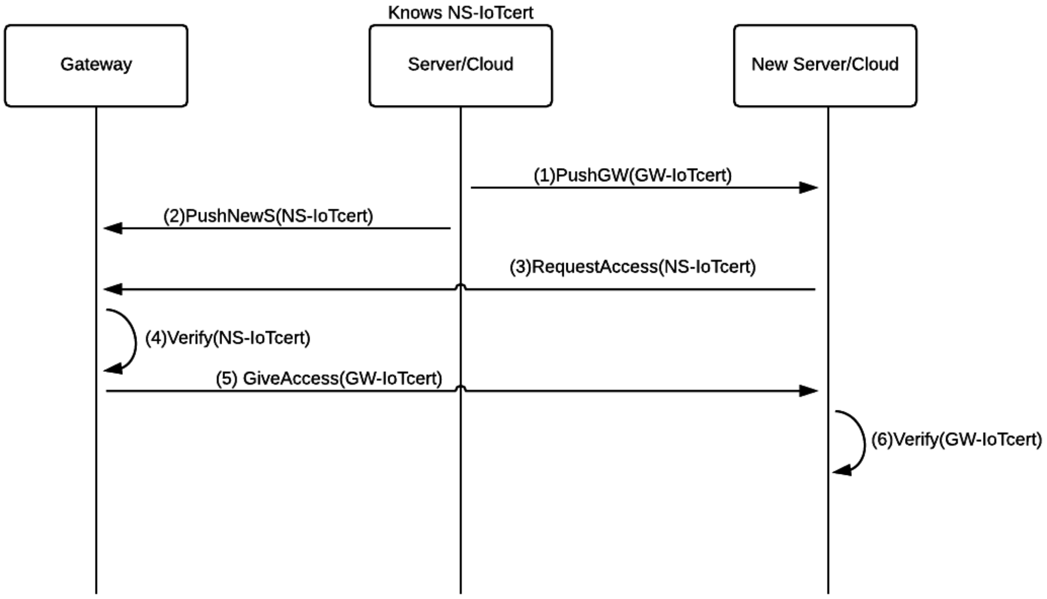


Figure 4—4: Access control

4.2.6 Conclusion

For this proposed method to be applicable there a need for adoption of some widely used standards, for example what type of communication protocol or key and crypto algorithm should be used (e.g. Rivest-Shamir-Adleman (RSA), Elliptic curve cryptography (ECC), or Diffie-Helman). This is just a general proposition and it will need extensive testing and analysing before actual implementation. The purpose of this proposal was to show that together with HIS and asymmetric cryptography, there could be a way of securing the identities (keys) and authentication in an IoT-environment, primarily against physical threats.

4.3 IoT Environment and case scenarios

This section examines three different scenarios were a threat arises in each of the three parts that are potential targets for an attacker: communication, IoT-device, and gateway. With the background information given in Chapter 2 and results from the interview study (see Chapter 4.1) it should be possible to present ideas functional solutions for the scenarios discussed below.

The environment we have chosen to analyse is shown in Figure 4—5. This IoT environment has been chosen because we have identified that a gateway is commonly used to interconnect constrained devices with the internet. The company and the device exchange information with each other, but we do not consider which peer initiates the communication.

The device is constrained in such a manner that it is less powerful than the gateway in regards to for example, CPU performance or memory space. The gateway is a powerful device which interconnects the 6LoWPAN network to the Internet. Section 2.8 explained that 6LoWPAN is a wireless network designed for constrained devices. The Internet is a global network where packets are automatically routed to the defined destination address. The gateway is fully capable of routing and compressing/decompressing IP packets that arrive either from the 6LoWPAN network or the Internet [70, 83, 84]. The traffic between the Server and the Gateway is IPv6. The Device and Gateway are also using IPv6 in a 6LoWPAN according to IEEE 802.15.4 standards [65, 66].

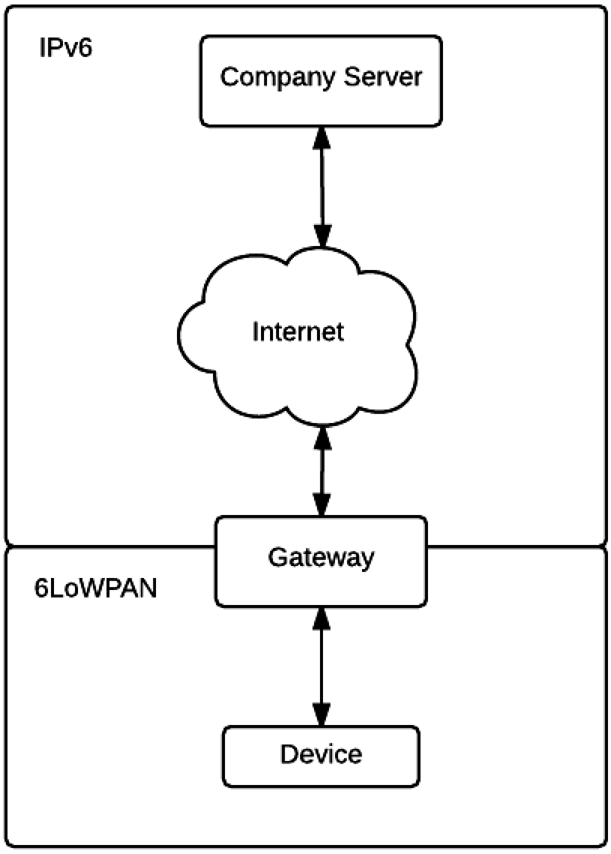


Figure 4—5: IoT Environment

As previously explained, three cases will be analysed in this IoT environment. Section 4.3.3 will discuss some of the aspects of a man-in- the-middle attack and eavesdropping on a communication level. These cases each have more detailed description in the corresponding subsection. The analysis of these scenarios is based on the analyses process described in Section 2.1.5 (with some slight modifications). In all cases the asset is information and the threat agent is an unauthorized person who tries to exploit the system to access the asset. The security components that will be analysed and identified are the threats, vulnerabilities, and risks. After all of these components have been identified, we will propose suitable safeguard(s). We choose to partly ignore the component's "Exposure" because it required more specifications about the information's purpose and value.

4.3.1 False gateway

As mentioned in Section 2.2.1.3, threats with a physical impact on an environment will be more common when an external attacker for example either modifies the existing gateway or replaces it with a clone (i.e., an exact copy as seen by the device and server). This could enable an attacker to see, modify, falsify, or redirect data sent from or to our devices. This depends on the fact that devices and gateways will be more easily accessible outside of a secure environment, which usually gives rise to physical tampering on some level.

The vulnerability could be a lack of data protection or because the authentication between nodes was non-existing (or a combination of both). As mentioned before, the first step is to ensure that a given device that is sending you data are your own device and that the data are unmodified between the device and the destination (see Section 4.3.3).

Even if you protect your data by encryption, an attacker could potentially retrieve the key and use it to create fake data messages to the gateway. For example, a pre-shared or exchanged symmetric key for data encryption could be stolen from a genuine device and transferred to another device (if not protected properly). If this other device does not need to be authenticated against the environment, it could send whatever data it wants as long as it has the correct key. Making the device non-tamperable in such a way that keys cannot be retrieved is an important step to securing an IoT-environment.

A solution to this problem would be the method proposed in chapter 4.2. By using the authentication method in chapter 4.2.2 the identity of the deployed gateway gets verified by authenticating it. Since we have confidence that the private key of the device is secured using HIS we have a bigger trust that the device is what it purports to be.

If an attacker would try creating a clone (exact copy) of the gateway the PUF would generate different PUF data and thereby generate a different private key in the key extractor function. The private key would not match the public key of the known identity and this would result in a gateway unable to authenticate itself against the Server/Cloud.

One disadvantage worth discussing with HIS is the possibility of software intrusion. Since the keys are obtainable by applying the activation code to the key extraction function, they might be retrievable provided that it is possible to do so from the any layer in the IoT-reference model (see Figure 1—2). For example, if the attacker applies modified software that can gain access to these activation codes and also apply them to the function, the output would be the private key, since it is still the same device (and thereby the same characteristics which generate correct PUF data). A solution to this would be to also authenticate the software running on the device.

4.3.2 False IoT-device

When it comes to IoT-devices the same concept explained in the previous subsection can be applied to secure the deployment and use of new devices. By authenticating the device against a gateway or server environment (see chapter 4.2.3) trust is created and reliable information can be sent or received between nodes. Since the device also uses HIS, the trust is amplified by knowing the private key is secure.

As explained in Section 4.2.4, revocation of compromised devices is easily done by revoking the certificate in the gateway which means that the device will not be able to authenticate against the gateway. The gateway could also have a limitation on failed authentication attempts which will limit the number of devices trying to authenticate against it. For example, after three failed attempts the public key could be temporarily or permanently banned.

4.3.3 Insecure communication

In this scenario the term insecure communication means that unprotected information is flowing between the server and client. This information is visible to anyone who is connected to the network.

A man-in-the-middle attack occurs when an attacker intercepts traffic between two peers, which an attacker either modifies or monitors [85]. Taken into account that the 6LoWPAN is a wireless network eavesdropping may also become relevant if the traffic is unprotected. In our specific case there are two possible areas where an attacker could position themselves. The attacker could be positioned somewhere on the internet as an intermediate node where it monitors or modifies information flowing between the server and gateway. The other option is the wireless network where the attacker eavesdrops traffic between gateway and device, see Figure 4—6. Monitoring (eavesdropping) and modifying are seen as two different attacks.

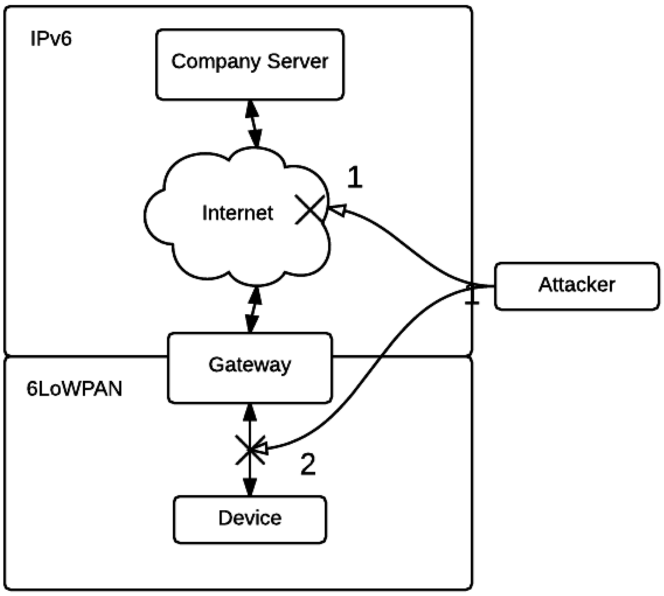


Figure 4—6: Man in the middle scenario

In this case, two threats are identified, the threat of an unauthorized person simply monitoring information and the threat of an unauthorized person modifying information. These threats exploit the vulnerability of an unprotected traffic flow and lack of authentication since the attacker can intercept/eavesdrop information that is traveling from the server to the device and vice versa. The vulnerabilities of an unprotected traffic flow leads to the risk of an attacker monitoring and modifying this information. The information monitored may be classed as confidential which leads to its confidentiality being broken. Modifying the information may lead to broken integrity and possibly broken confidentiality as well.

Preventing eavesdropping of information that is sent between gateway and device (point number 2 in Figure 4—6) could be solved by using the standardised link-layer security for IEEE 802.15.4 networks. The mandated security method offers encryption and message authentication for integrity and confidentiality by using shared keys between two or more peers [65]. Link-layer security works well for the actual IEEE 802.15.4 network, but does not provide any protection outside of the network when the packet is forwarded over the actual Internet. Link-layer security works on a hop-by-hop basis meaning that incoming packages are decrypted and then encrypted again before transmission to the next node [68, 65]. Link layer-security is a viable option against eavesdropping in the 6LoWPAN but does not remove the need of a trustworthy gateway since it has to decrypt and encrypt ingoing and outgoing information. If the gateway is compromised it would jeopardize all traffic.

Since the gateway does not provide any services than forwarding and decompressing/compressing a desired solution would be use end-to-end protection between server and device. This would also remove the possibility of a man-in-the-middle node monitoring or modifying information traveling between gateway and server.

Two possible end-to-end solutions have been explained in Chapter 2: DTLS and IPsec. Both of these protocols have been tested for 6LoWPAN with successful header compression for both protocols while using pre-shared keys. Since the devices used in an IEEE 802.15.4 network are constrained a further analysis has to be made in regards what type of security needs to be implemented. In this scenario there are two categories of data that are relevant. One category is when data is not confidential, but is in need of integrity protection. The other category is when data needs to be confidential and integrity protected. The latter case is considered because it does not make sense to ensure data confidentiality *without* ensuring its integrity. These two categories might require different components. This way we might avoid using more security functions than needed.

4.3.3.1 Data integrity

In this case it does not matter if an attacker is monitoring the data as long as there are no alterations to the data. In this case the data does not have to be encrypted, although its integrity has to be ensured. Both IPsec and DLTS provide this by signing messages with a MAC which ensures data integrity. If pre-shared keys are used, DTLS offers three different kinds of cipher suites which provide a hash function without the need for encryption [86]. As for IPsec this would mean using the AH header. Compared to DTLS, AH ensures data origin in addition to integrity protection which may be desirable in some cases. Because it also protects the integrity of the IP header as explained in Section 2.5. ESP in IPsec can be used in a similar way to DTLS since it also has the possibility of providing data integrity without encryption [87].

4.3.3.2 Confidentiality and data integrity

If the data is confidential one most likely wishes to ensure its integrity as well. In this case DTLS works in a normal fashion by first computing a MAC and then encrypting the data. In IPsec the logical choice would be to use ESP since it offers confidentiality and optionally integrity. The difference in the process when using DTLS is that ESP reverse the order of process by first encrypting the data and then computing a MAC. For a constrained device this might be relevant because the device can immediately discard a tampered message if the authentication fails. As for DTLS the device would first have to decrypt the message in order to notice any alterations.

4.3.3.3 Comparing IPsec and DLTS

Both protocols can offer the needed level of security in order to secure the environment in Figure 4—6. As explained in Chapter 2, DTLS provides its own reliability on top of UDP. DTLS is implemented between the application layer and the transport layer. DTLS requires the use of applications or operating systems which support DTLS. This may put a constraint on which applications can be used in a system. This creates dependence on layers above and below DTLS. IPsec is more generally applicable in that it does not matter which protocols are used on the transport and application layer. IPsec requires that each endpoints in order to process the packets. Choosing between IPsec and DTLS may be difficult because it requires further research on both protocols and comparing these on the same hardware. Another problem which both these protocols face in 6LoWPAN is the lack asymmetric authentication since both implementations are based on pre-shared keys. The use of a shared key works well in this environment, but adding additional endpoints would mean a lot of manual configuration which is not ideal in regards to scalability.

An interesting notation is that a draft has been made for a minimal IKE version which is adapted for constrained devices. Although this draft purposes a version of IKE, it has left many functions out [88].

It would also be interesting to combine the trust model in Chapter 4.2 with either DTLS or IPsec since the trust model reduces the possibility of a false gateway or device trying to communicate with the server.

DTLS and IPsec together with pre-shared keys are the proposed solutions with IPsec as a personal preference, since it does not rely on what is running on the application or transport layer thus making it more flexible. It would probably be possible to implement both protocols at the same time, but this would be undesirable for 6LoWPAN since the purpose of header compression is to make the packets smaller. Unnecessarily stacking protocols would mean less space for sending actual data.

Notable is that we do not consider the energy use, the difficulty of implementing, or the computational cost for any of these protocols; thus an implementation would require more specifications and testing.

4.3.4 Conclusion

The end-to-end authentication method presented in Chapter 4.2.2 combined with the IPsec solution in Chapter 4.3.3 could possibly secure the IoT-environment. The combined solution would require some additional functions in the trust model in order to enable key-exchanges and the establishment of security associations. An additional function could be the addition of a Diffie-Hellman key exchange during the authentication process in order to establish symmetric keys used for encryption and hashing in IPsec. This may mean an additional two messages in the authentication to ensure that keys were correctly exchanged.

The result of both methods would provide devices with a unique identity which could be used for authentication and key-exchanges in order to securely establish an IPsec connection.

5 Analysis

In this chapter, we discuss the results from the interviews and scenarios in Chapter 4 and analyse these in relation to the information provided in Chapters 1 and 2.

One of the major parts for IoT is the provision of information as a basis for decision-making. Whether the purpose is to provide information to a smaller business system or larger so-called system of systems, the decision-making in the end revolves around information. With systems composed of multiple parts such as a central system together multiple decentralized subsystems providing information there will be a crucial need for trust among these systems. Establishing trust among systems might involve a long chain of trusts. For example, two larger third-party systems providing information to each other will need to trust each other. If a third-party system consists of multiple sub-systems it will create an even long chain of trust among the third-party's sub-systems. You will have to trust that the third-party has a secure chain of trust inside their system. Depending on the situation these chains may become very long and complicated. As mentioned in chapter 2.4.2, this can be solved with certificate chain reduction [56].

One of the conclusions drawn from the background is that it is necessary to identify de-perimeterised devices and do so in a secure way by creating trust chains. The method described includes asymmetric cryptography were the private key is reconstructed (with HIS) each time it needs to be used and is thereby protected from physical threats that involves tampering and cloning of a device. HIS is a tested concept that is used today in certain solutions and can thereby be seen as somewhat secure method to use [89]. As proposed in this thesis, this could very well be an important method to use for securing the IoT. It is important to know that this method only identifies the hardware of the device since the key is linked to the hardware characteristics. To further secure the whole device, authentication of the application or identifying changes in the software running on the device is necessary. This could be done with tamper-proofing, for example if an attacker tries to remove a watermark from software [90].

In order to achieve trust among systems an important part will be to secure them. We believe that the approach to securing these systems will continue to rely on threat and risk analyses, but the solutions will consist of many different kinds of security architectures.

Securing IoT-environments will be a difficult task since there will be many different use scenarios and every scenario consisting of different kind of devices. Each security solution will look different from the other since these systems may contain entities which are constrained in different ways. As stated in Interview number 3, one needs to know where the system is vulnerable and in order to identify these vulnerabilities it is vital to know both the hardware and software specifications. These specifications aid the identification of vulnerabilities because systems limits are defined, but it also helps find applicable solutions. A solution for IoT will either have to be very general or very specific because defining solutions on a level in between will leave many unsolved problems. A positive aspect of a general solution is that it may provide guidelines in the process of designing a perfect solution.

The cases discussed in Section 4.3 could most likely contribute to securing an IoT-environment, although a realistic IoT-environment will most likely be exposed to more threats that need to be addressed in order to secure it to a desired level. A threat and risk analysis will identify many different kinds of threats, but this does not exclude the possibility of one solution countering multiple threats. For example, the implementation of IPsec offers many services, such as access control, data integrity, and data confidentiality - which all may address different threats.

One of the characteristic of IoT is its expected “enormous scale” as there will be many interconnected devices. In interview number 1 it was stated that a goal of IoT was to create so-called systems of systems. In relation to this the question of operating at an “enormous scale” means that each of these systems needs a thorough security analysis. A security analysis or a threat and risk analysis will not only include software security because if a systems is de-perimeterised and devices are outside the perimeter of a secure company environment, then physical threats will become more and more relevant. A desired level of security has to be found which provides enough safety without affecting the functionality too much.

As stated in Section 1.2.1, security may be implemented at every layer in an IoT device, but in some cases multiple security implementations might not be possible. Stacking security protocols on a constrained device might provide security in multiple levels (i.e. IP, transport and application-layer) , but in 6LoWPAN it might limit the amount of data that can be sent thus negatively affecting its functionality.

The solutions in Section 4.3 are well designed for those specific scenarios, but are very limited in the sense of scalability. For example, the possible solutions today are not scalable enough since both DTLS and IPsec in 6LoWPAN use pre-shared keys which require manual configuration. Thus adding additional devices or servers would require an administrator to configure both ends in the communication flow (server or device). Security solutions will have to adapt to these constrained devices meaning that solutions might have to change in order to be implementable in more cases than normal (e.g. home-PC to server communication). That is why it is interesting seeing adaptation to protocols such as DTLS, IPv6, and IPsec so these can be used in IEEE 802.15.4 network without having a too large effect on devices functionality.

Securing an IoT-environment will revolve around the classification of information. There will have to be a balance between the value of information and the security capabilities of a device. The challenge may not be securing a central system (e.g. server) as this will not be more difficult than it is today instead, the challenge may lie in securing the subsystems (e.g. constrained device) where the limits of the subsystems will restrict the security capabilities. There are defined standards for securing normal systems and internet-communication that exist today, but these standards must become more lightweight (e.g. header compression for 6LoWPAN) in order to adapt to these new security challenges concerning constrained devices.

6 Conclusions and Future work

This chapter contains the conclusions drawn from working with the concept and the actual results of the thesis research together with future work that suggest development and implementation of a unified authentication method.

6.1 Conclusions

The goals (stated in Section 1.7) of this thesis were to introduce the reader to the IoT concept, identify security challenges that need to be addressed in order to secure IoT, propose a solution to them, and to do an interview study in regards to information security in IoT. The first goal resulted in a rather extensive introduction (see Chapter 1) to the thesis which took a little longer than initially expected, but was of big help when working further to identify the security challenges (see Chapter 2) and proposing our solution (see Section 4.2) to the identity and authentication problem and secure communication (see Section 4.3.3). Unfortunately the results of interview number 3 were not what we wished for since our IoT-environment was not specific enough, but none the less provides valuable insight about security designs. The final goal was not achieved as planned because of miscommunication, but the results helped us in the process of writing this thesis.

IoT is a very interesting concept which creates many new possibilities in form of services and inventions. IoT is an enormously extensive concept that only has very general requirement postures or very specific solutions depending on how specifically you look at it.

There is a lot of research in many different areas involving IoT. Many different researchers have proposed many different kinds of adaptations to protocols and authentication methods for IoT which makes it very difficult to identify the best solution. Therefore, there is grave need of structured guidelines in the form of standardisation in order to interconnect all kinds of devices, protocols, applications, etc. Developing standards or solutions needs to come with open source protocols and methods in order to attract wide acceptance and use. By trying to give an understanding of how such a standard should be developed and what requirements are needed, we hope that we have helped layer a foundation for further studies in the area.

A major part of the research concerned IP-based communication based on IPv6. The use of IPv6 will probably be crucial for finding and addressing unique devices on the Internet since the number of available addresses will enable every device to have its own IP address.

When working with such an extensive concept as IoT, it is both difficult and time consuming to fully understand. This lead to us delimiting our work at a rather late state and therefore we had the problem of deciding on which area to focus. If this report had to be done again, the delimitation would have been done earlier by spending less time trying to understand the actual IoT concept. For research which is limited in time it is important to delimit the research focus. This becomes more important if the goal is to provide an extensive explanation. If the goal of a project is to develop some kind of software it is easier to delimit the work because there is a clear goal from start to finish. Research which mostly resolves around literature studies becomes important to define a goal that is very specific, otherwise it easy to spin off because there is a lot of information available.

Although there is a lot more work to be done before actually implementing the method proposed for identification, authentication and trust; this proposal can be used as a basis for further development in security methods in correlation to IoT.

As the conclusion in Section 4.2.6 states, the proposed authentication method needs extensive (more than half a bachelor thesis project) analysis, for example how to deal with replay attacks, identity management on the device, which protocols to be used for communication, which cryptographic algorithms and keys to use, etc. These properties need to be carefully chosen and implemented if a standard is to be developed, which is put off to future work.

6.2 Limitations

What limited our efforts the most was limited amount of time for this thesis project. IoT is such an extensive concept that it was impossible to research everything during 10 weeks. Another limitation was the lack of knowledge about certain protocols, such as Bluetooth and Zigbee. This required us to ignore the possibilities of these protocols on a technical level because we did not have enough time to acquire the knowledge needed to fully understand all of the different protocols. Information security was a new area for us and this required additional work in order to understand the concept and apply it to our research.

6.3 Future work

It would be of interest to find out if the identification and authentication method proposed is indeed applicable in a typical IoT-device and environment. That is, is it possible to store secrets keys this way on constrained devices and to distribute certificates and rights? By creating a unified identity and authentication method for all devices connected to the internet, the “communicate with any Thing” requirement would be fulfilled. This is desired in order to connect different systems, independent on what information that needs to be sent. This creates dynamics in different environments by being able to interconnect any Thing and thereby exchange information securely.

To further realise our proposal, the following work is required:

- Secret key storage - Implement a PUF on a constrained device and try to generate keys

  Use these keys to authenticate the application layer (e.g. sign application certificates with this hardware key) or implement some tamper-proofing in the software.

- Identify vulnerabilities (e.g. replay attacks) in the proposed authentication method and find solutions to them before implementation

  Perhaps combine with IPsec?

- Further research and implementation of key exchanges together with security protocols for IP-communication in constrained networks.
- Research and compare IP-connections for Bluetooth and ZigBee [91, 92]
- Future work should consider IoT devices that cannot store a certificate and the certificate of a CA.
- Compare DTLS and IPsec with a " IKEv2 based lightweight secure data communication” [93]

Hopefully our work can provide an understanding for IoT and guidelines for additional research.

6.4 Required reflections

Generally the concept of IoT offers many possibilities for both environmental and economic sustainability by making different processes more efficient. For example, using sensors to make energy consumption more efficient might enable a person or company to save money by using less energy. Additionally, using less energy may contribute to lowering energy production at for example a power plant. Nothing states that IoT has to be or will be used this way but the technology will at least offer the possibility. The one implementing an IoT product in the future will be responsible of how this may or may not affect the aspects of sustainability.

With more sensors and devices monitoring our everyday life and environment, the question of how much and how specific the data collected should be arises. Where is the boundary between information gathering and the violation of personal integrity?

Compared to a device not using HIS, this method will make it pretty much unsustainable for an attacker trying to retrieve the private key on a device, gateway or server to further exploit the system. Even if a device gets compromised or stolen, it could be physically identifiable with DNA. Depending on the cost of the device, it might just be economically viable to replace it with a new one instead of trying to recover the old one, since it probably needs to get enrolled again.

References

[1] K. Ashton, "That 'Internet of Things'," 22 June 2009. [Online]. Available: http://www.rfidjournal.com/articles/view?4986. [Accessed 02 April 2015].

[2] Gartner, "Gartner Says the Internet of Things Installed Base Will Grow to 26 Billion Units By 2020," Gartner, 12 December 2013. [Online]. Available: http://www.gartner.com/newsroom/id/2636073. [Accessed 02 April 2015].

[3] Cisco, ""Connections Counter The Internet of Everything in Motion, " Cisco's The Network," Cisco, 29 July 2013. [Online]. Available: http://newsroom.cisco.com/feature/1208342/Connections-Counter-The-Internet-of-Everything-in-Motio_2. [Accessed 29 April 2015].

[4] EMC & IDC, "The Internet of Things: Data from Embedded Systems Will Account for 10% of the Digital Universe by 2020 | The Digital Universe of Opportunities: Rich Data and Increasing Value of the Internet of Things," EMC & IDC, April 2014. [Online]. Available: http://www.emc.com/leadership/digital-universe/2014iview/internet-of-things.htm. [Accessed 02 April 2015].

[5] G. E. Moore, "Cramming More Components onto Integrated Circuits," [Online]. Available: http://www.cs.utexas.edu/~fussell/courses/cs352h/papers/moore.pdf. [Accessed 24 April 2015].

[6] Intel, "Intel Chips," 7 December 2013. [Online]. Available: http://www.intel.com/content/dam/www/public/us/en/documents/corporate-information/history-intel-chips-timeline-poster.pdf. [Accessed 24 April 2015].

[7] J. G. Koomey, S. Berard, M. Sanchez and H. Wong, "Assesing Trends In The Elecritcal Efficiency Computation Over Time," 2009.

[8] K. Ashton, "Kevin Ashton. The Internet of Things. Seoul, June 19, 2014 - YouTube," 19 June 2014. [Online]. Available: https://www.youtube.com/watch?v=xSYkp8_Dn2E. [Accessed 15 April 2015].

[9] Postscapes, "Internet of Things Definition- Postscapes," Postscapes, [Online]. Available: http://postscapes.com/internet-of-things-definition. [Accessed 13 April 2015].

[10] O. Vermesan and P. Friess, Internet of Things - From Research and Innovation to Market Deployment, Aalborg Ø: River Publishers, 2014.

[11] ITU Telecommunication Standardization Sector, "ITU-T Recommendation database," 2012. [Online]. Available: http://handle.itu.int/11.1002/1000/11559-en?locatt=format:pdf&auth. [Accessed 13 April 2015].

[12] Cisco Systems, "Internet of Things (IoT) - Cisco Systems," [Online]. Available: http://www.cisco.com/web/solutions/trends/iot/overview.html. [Accessed 14 April 2015].

[13] IBM, "The Internet of Things - YouTube," IBM Social Media, 15 March 2010. [Online]. Available: https://www.youtube.com/watch?v=sfEbMV295Kk. [Accessed 14 April 2015].

[14] J. Barrett, "The Internet of Things Dr. John Barrett at TEDxCIT - YouTube," TEDx Talks, 5 October 2012. [Online]. Available: https://www.youtube.com/watch?v=QaTIt1C5R-M. [Accessed 14 April 2015].

[15] "CASAGRAS an EU Framework 7 Project (Coordination and Support Action for

Global RFID-related Activities and Standardisation),” [Online]. Available: http://grifs-project.uniweb.be/data/File/CASAGRAS%20FinalReport%20(2).pdf. [Accessed 15 April 2015].

[16] Postscapes, "Internet of Things Examples - Postscapes," [Online]. Available: http://postscapes.com/internet-of-things-examples/. [Accessed 17 April 2015].

[17] Bigbelly, “Bigbelly,” 2015. [Online]. Available: http://bigbelly.com/. [Accessed 4 May 2015].

[18] I. Wigmore, "What is Internet of Things (IoT) - Definition from WhatIs.com," [Online]. Available: http://whatis.techtarget.com/definition/Internet-of-Things. [Accessed 10 April 2015].

[19] D. A. Gustafson, “Artificial intelligence - AccessScience from McGraw-Hill Education,” AccessScience, 2014. [Online]. Available: http://www.accessscience.com.focus.lib.kth.se/content/artificial-intelligence/053300. [Accessed 4 May 2015].

[20] L. Guixiong, X. Jialong and H. Xiaobin, “Internet of Things Sensor Node Information Scheduling Model and Energy Saving Strategy,” *Advanced Materials Research*, vol. 773, pp. 215-220, 2013.

[21] X. Guo, Z. Wang and L. Zhao, “Intelligent Industrial Park based on Internet of Things,” *Advanced Materials Research*, vol. 722, pp. 486-490, 2013.

[22] Z. Peng, Z. Jingling and L. Qing, “Message Oriented Middleware Data Processing Model In Internet of Things,” in *Internation Conference on Computer Science and Network Technology*, Changchun, 2012.

[23] Federal Trade Commission (FTC), "Internet of Things Privacy & Security in a Connected World," 2015. [Online]. Available: https://www.ftc.gov/system/files/documents/reports/federal-trade-commission-staff-report-november-2013-workshop-entitled-internet-things-privacy/150127iotrpt.pdf. [Accessed 13 April 2015].

[24] BMW, “BMW ConnectedDrive - Overview,” BMW, [Online]. Available: http://www.bmw.com/com/en/insights/technology/connecteddrive/2013/index.html. [Accessed 17 April 2015].

[25] ADAC, “ADAC Info - Fahrerassistenzsysteme,” [Online]. Available: https://www.adac.de/infotestrat/technik-und-zubehoer/fahrerassistenzsysteme/sicherheitsluecken.aspx. [Accessed 17 April 2015].

[26] C. G. Weissman, "We Asked Executives About The Internet Of Things And Their Answers Reveal That Security Remains A Huge Concern," Business Insider, [Online]. Available: http://www.businessinsider.in/We-Asked-Executives-About-The-Internet-Of-Things-And-Their-Answers-Reveal-That-Security-Remains-A-Huge-Concern/articleshow/45959921.cms. [Accessed 10 April 2015].

[27] Jericho Forum, "Jericho Forum Commandments," 2007. [Online]. Available: https://collaboration.opengroup.org/jericho/commandments_v1.2.pdf. [Accessed 13 April 2015].

[28] ITU-T, "A Handbook on Internet Protocol (IP)-based Networks and Related Topics and Issues," 2005. [Online]. Available: http://www.itu.int/ITU-T/special-projects/ip-policy/final/IPPolicyHandbook-E.pdf. [Accessed 13 April 2015].

[29] Information Sciences Institute, "Internet Protocol," RFC 791, September 1981. [Online]. Available: https://www.ietf.org/rfc/rfc791.txt. [Accessed 13 April

2015].

[30] S. Deering and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specfication," RFC 2460 (Draft Standard), December 1998. [Online]. Available: https://www.ietf.org/rfc/rfc2460.txt. [Accessed 10 April 2015].

[31] KTH, “KTH IK121X Examensarbete inom kommunikationssystem, grundnivå 15,0 hp,” [Online]. Available: https://www.kth.se/student/kurser/kurs/IK121X. [Accessed 20 April 2015].

[32] LII, “44 U.S. Code § 3532 - Definitions LII - Legal Information Institute,” [Online]. Available: https://www.law.cornell.edu/uscode/text/44/3532. [Accessed 21 April 2015].

[33] J. Andress, The Basics of Information Security, Syngress, 2011.

[34] S. Bosworth, E. Whyne and M. Kabay, Computer Security Handbook, Fifth Edition, John Wiley & Sons, 2009.

[35] G. Stoneburner, C. Hayden and A. Feringa, “Engineering Principles for Information Technology Security (A Baseline for Achieving Security),” National Institute of Standards and Technology, Gaithersburg, 2004.

[36] OWASP, “OWASP Internet of Things Top Ten Project - OWASP,” 5 May 2015. [Online]. Available: https://www.owasp.org/index.php/OWASP_Internet_of_Things_Top_Ten_Project#tab=OWASP_Internet_of_Things_Top_10_for_2014. [Accessed 15 May 2015].

[37] M. McDowell, “Understanding Denial-of-Service Attacks US-CERT,” United States Computer Emergency Readiness Team, 2013. [Online]. Available: https://www.us-cert.gov/ncas/tips/ST04-015. [Accessed 25 May 2015].

[38] IEC, “IEC - Electromagnetic compatibility - EMC explained EMC and the IEC,” 2015. [Online]. Available: http://www.iec.ch/emc/explained/. [Accessed 25 May 2015].

[39] OWASP, “Man-in-the-middle attack - OWASP,” OWASP, 8 April 2014. [Online]. Available: https://www.owasp.org/index.php/Man-in-the-middle_attack. [Accessed 25 May 2015].

[40] OWASP, “Top 10 2014-I5 Privacy Concerns - OWASP,” OWASP, 2 April 2015. [Online]. Available: https://www.owasp.org/index.php/Top_10_2014-I5_Privacy_Concerns. [Accessed 25 May 2015].

[41] P. Legonkov and V. Prokopov, “Small Cell Wireless Backhaul in Mobile Heterogeneous,” July 2015. [Online]. Available: http://urn.kb.se/resolve?urn=urn%3Anbn%3Ase%3Akth%3Adiva-99010. [Accessed 8 June 2015].

[42] I. Friese, J. Stollman and S. Shorter, “Concepts of Identity within the Internet of Things - DG - Identities of Things - Kantara Initiative,” 28 April 2014. [Online]. Available: http://kantarainitiative.org/confluence/display/IDoT/Concepts+of+Identity+within+the+Internet+of+Things. [Accessed 22 May 2015].

[43] H. Handshuh, G.-J. Schrijen and P. Tuyls, “Hardware Intrinsic Security from Physically Unclonable Functions,” 2011. [Online]. Available: http://www.springer.com/cda/content/document/cda_downloaddocument/9783642144516-c1.pdf?SGWID=0-0-45-1014345-p174025367. [Accessed 26 May 2015].

[44] Gemalto NV, “Hardware Security Modules (HSMs) SafeNet Encryption & Key Security,” 2015. [Online]. Available: http://www.safenet-inc.com/data-

encryption/hardware-security-modules-hsms/. [Accessed 25 May 2015].

[45] J. Schlyter, “HSM - Hardware Security Module,” 2009. [Online]. Available: https://www.iis.se/docs/hsm-20090529.pdf. [Accessed 25 05 2015].

[46] B. Gassend, D. Clarke, D. Lim, M. v. Dijk and S. Devadas, “Identification and Authentication of Integrated Circuits,” Computer Science and Artificial Inteligence Laboratory - Massachusetts Institute of Technology, June 2003. [Online]. Available: http://csg.csail.mit.edu/pubs/memos/Memo-466/memo-466.pdf. [Accessed 22 May 2015].

[47] S. Devadas, D. Clarke, B. Gassend, D. Lim, J. Lee and M. v. Dijk, “Physical Unclonable Functions and Applications,” [Online]. Available: http://people.csail.mit.edu/rudolph/Teaching/Lectures/Security/Lecture-Security-PUFs-2.pdf. [Accessed 25 May 2015].

[48] G.-J. Schrijen and V. v. d. Leest, “Comparative analysis of SRAM memories used as PUF primitives,” [Online]. Available: http://www.intrinsic-id.com/wp-content/uploads/2014/09/SRAM-memories.pdf. [Accessed 27 May 2015].

[49] NXP Semconductors N.V., “PUF - Physically Unclonable Functions: Protecting next-generation Smart Card ICs with SRAM-based PUFs,” 2013. [Online]. Available: http://www.nxp.com/documents/other/75017366.pdf. [Accessed 27 May 2015].

[50] Microsoft, “Public Key Infrastructure (Windows),” [Online]. Available: https://msdn.microsoft.com/en-us/library/windows/desktop/bb427432(v=vs.85).aspx. [Accessed 31 May 2015].

[51] Microsoft, “X.509 Public Key Certificates (Windows),” [Online]. Available: https://msdn.microsoft.com/en-us/library/windows/desktop/bb540819(v=vs.85).aspx. [Accessed 31 05 2015].

[52] Microsoft, “Trust Models (Windows),” [Online]. Available: https://msdn.microsoft.com/en-us/library/windows/desktop/bb540815(v=vs.85).aspx. [Accessed 31 May 2015].

[53] Microsoft, “Certificate Hierarchy (Windows),” [Online]. Available: https://msdn.microsoft.com/en-us/library/windows/desktop/bb931353(v=vs.85).aspx. [Accessed 31 May 2015].

[54] Microsoft, “Cross Certification (Windows),” [Online]. Available: https://msdn.microsoft.com/en-us/library/windows/desktop/bb540800(v=vs.85).aspx. [Accessed 31 May 2015].

[55] Microsoft, “Certificate Chain (Windows),” [Online]. Available: https://msdn.microsoft.com/en-us/library/windows/desktop/bb540794(v=vs.85).aspx. [Accessed 31 May 2015].

[56] Y. Kortesniemi, “SPKI Performance and Certificate Chain Reduction,” [Online]. Available: http://subs.emis.de/LNI/Proceedings/Proceedings19/GI-Proceedings.19-71.pdf. [Accessed 8 June 2015].

[57] S.Kent; and K.Seo, “Security Architecture for the Internet Protocol (IPsec),” RFC 4301 (Draft Standard), December 2005. [Online]. Available: https://tools.ietf.org/html/rfc4301. [Accessed 29 May 2015].

[58] C.Kaufman, “Internet Key Exchange Protocol Version 2 (IKEv2),” RFC 5996 (Draft

Standard), September 2010. [Online]. Available: http://tools.ietf.org/html/rfc5996/. [Accessed 29 May 2015].

[59] S.Sakane, K.Kamada, M.Thomas and J.Vilhuber, “Kerberized Internet Negotiation of Keys (KINK),” RFC4430, March 2006. [Online]. Available: http://tools.ietf.org/html/rfc4430. [Accessed 29 May 2015].

[60] T.Dierks and E.Rescorla, “The Transport Layer Security (TLS) Protocol Version 1.2,” RFC 5246 (Draft Standard), August 2008. [Online]. Available: https://tools.ietf.org/html/rfc5246. [Accessed 29 May 2015].

[61] S. Vandeven, “SSL/TLS: What's Under the Hood?,” SANS Institute InfoSec Reading Room, 13 August 2013. [Online]. Available: http://www.sans.org/reading-room/whitepapers/authentication/ssl-tls-hood-34297. [Accessed 29 May 2015].

[62] P. Sjödin, *SSL/TLS - Lecture in the course "IK2206 Internet Security and Privacy (Fall 2014)*, Stockholm: KTH, 2014.

[63] Information Sciences Institute University of Southern California, “Transmission Control Protocol,” RFC 793 (Draft Standard), September 1981. [Online]. Available: http://tools.ietf.org/html/rfc793. [Accessed 9 June 2015].

[64] E.Rescorla and N.Modadugu, “Datagram Transport Layer Security (DTLS) Verion 1.2,” RFC 6347 (Draft Standard), January 2012. [Online]. Available: https://tools.ietf.org/html/rfc6347. [Accessed 30 May 2015].

[65] IEEE Computer Society, IEEE Standard for Local and metropolitan area networks — Part 15.4 (802.15.4): Low-Rate Wireless Personal Area Networks (LR-WPANs), Unted States of America: The Institute of Electrical and Electronics Engineers, 2011.

[66] “IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs): Overview, Assumptions, Problem Statement, and Goals,” RFC 4919 (Informational Draft), August 2007. [Online]. Available: https://tools.ietf.org/html/rfc4919. [Accessed 30 May 2015].

[67] Z. Shelby and C. Bormann, 6LoWPAN : The Wireless Embedded Internet, John Wiley & Sons Ltd, 2009.

[68] S.Raza, S.Duquennoy, J.Höglund, U.Roedig and T.Voigt, “Secure Communication for the Internet of Things - A Comparison of Link-Layer Security and IPsec for 6LoWPAN,” *Security and Communication Networks,* vol. 7, no. 12, pp. 2654-2668, 2014.

[69] S. Raza, S. Duquennoy, T. Chung, D. Yazar, T. Voigt and U. Roedig, “Securing Communication in 6LoWPAN with Compressed IPsec,” 11 August 2011. [Online]. Available: http://soda.swedish-ict.se/4183/1/raza11securing.pdf. [Accessed 30 May 2015].

[70] S. Raza, S. Duquennoy and G. Selander, “Compression of IPsec AH and ESP Headers for Constrained Environments,” Internet-Draft, 3 September 2013. [Online]. Available: https://tools.ietf.org/html/draft-raza-6lowpan-ipsec-01. [Accessed 30 May 2015].

[71] S.Raza, H.Shafagh and O.Dupont, “Compression of Record and Handshake Headers for Constrained Environments,” IETF - Internet Draft, 10 March 2014. [Online]. Available: https://tools.ietf.org/html/draft-raza-dice-compressed-dtls-00. [Accessed 30 May 2015].

[72] S.Raza, H.Shafagh, K.Hewage, R.Hummen and T.Voigt, “Lithe: Lightweight Secure CoAP for the Internet of Things,” *Sensors Journal, IEEE,* vol. 13, no. 10, pp. 3711-3720, 2013.

[73] Z. Shelby, K. Hartke and C. Bormann, “The Constrained Application Protocol (CoAP),” RFC 7252 (Standard Draft), June 2014. [Online]. Available: https://tools.ietf.org/html/rfc7252. [Accessed 30 May 2015].

[74] S. Raza, H. Shafagh, K. Hewage, R. Hummen and T. Voigt, “Lithe: Lightweight Secure CoAP for the Internet of Things,” *Sensors Journal, IEEE,* vol. 13, no. 10, pp. 3711-3720, 2013.

[75] C.Bormann, “Using CoAP with IPsec,” IETF - Information Draft, 6 December 2012. [Online]. Available: https://tools.ietf.org/html/draft-bormann-core-ipsec-for-coap-00. [Accessed 30 May 2015].

[76] L. Maqueda Ara, “Neighbor Discovery Proxy-Gateway for 6LoWPAN-based Wireless Sensor Networks : Design, Implementation, Analysis, and Evaluationm, Trita-ICT-EX-2011-221,” 21 December 2011. [Online]. Available: http://kth.diva-portal.org/smash/record.jsf?pid=diva2%3A470434&dswid=-5918. [Accessed 9 June 2015].

[77] J.Arkko, J.Kempf, B.Zill and P.Nikander, “SEcure Neighbor Discovery (SEND),” RFC 3971 (Draft Standard), March 2005. [Online]. Available: https://tools.ietf.org/html/rfc3971. [Accessed 9 June 2015].

[78] KTH, “Primo by Ex Libris,” KTH, [Online]. Available: http://kth-primo.hosted.exlibrisgroup.com/primo_library/libweb/action/search.do?vid=KTH. [Accessed 1 June 2015].

[79] Google, “Google,” [Online]. Available: https://www.google.se/. [Accessed 1 June 2015].

[80] General Electric, “Industrial Internet: Pushing the Boundaries of Minds and Machines,” 26 November 2012. [Online]. Available: http://www.ge.com/sites/default/files/Industrial_Internet.pdf. [Accessed 8 May 2015].

[81] D. Kushner, “The Real Story of Stuxnet,” IEEE Spectrum, 26 February 2013. [Online]. Available: http://spectrum.ieee.org/telecom/security/the-real-story-of-stuxnet. [Accessed 8 May 2015].

[82] M. Tesfaye, “Secure Reprogramming of a Network Connected Device : Securing programmable logic controllers,” 25 December 2012. [Online]. Available: http://kth.diva-portal.org/smash/record.jsf?pid=diva2%3A562977&dswid=-1262. [Accessed 9 June 2015].

[83] S. Raza, H. Shafagh and O. Dupont, “Compression of Record and Handshake Headers for Constrained Environments,” Internet Draft, 10 March 2014. [Online]. Available: https://tools.ietf.org/html/draft-raza-dice-compressed-dtls-00. [Accessed 28 May 2015].

[84] J. Hui and P. Thubert, “Compression Format for IPv6 Datagrams over IEEE 802.15.4-Based Networks,” RFC 4944 (Standard Draft), September 2011. [Online]. Available: https://tools.ietf.org/html/rfc6282. [Accessed 30 May 2015].

[85] Symantec Corporation, “Man-in-the-middle attack,” Symantec Corporation, [Online]. Available: http://www.symantec.com/security_response/glossary/define.jsp?letter=m&word=man-in-the-middle-attack. [Accessed 30 May 2015].

[86] Internet Assigned Numbers Authority, “Transport Layer Security (TLS) Parameters,” Internet Assigned Numbers Authority (iana), 12 May 2015. [Online]. Available: http://www.iana.org/assignments/tls-parameters/tls-parameters.xhtml. [Accessed 31 May 2015].

[87] R.Glenn and S.Kent, “The NULL Encryption Algorithm and Its Use With IPsec,” RFC

2410 (Draft Standard), November 1998. [Online]. Available: https://tools.ietf.org/html/rfc2410. [Accessed 31 May 2015].

[88] T.Kivinen, “Minimal IKEv2,” IETF (Informational Draft), 1 October 2012. [Online]. Available: https://tools.ietf.org/html/draft-kivinen-ipsecme-ikev2-minimal-01. [Accessed 31 May 2015].

[89] Intrinsic ID, “Intrinsic ID – Quiddikey®,” 2015. [Online]. Available: https://www.intrinsic-id.com/products/quiddikey/. [Accessed 2 June 2015].

[90] C. Thomborson, J. Nagra, R. Somaraju and C. He, “Tamper-proofing Software Watermarks,” [Online]. Available: http://crpit.com/confpapers/CRPITV32Thomborson.pdf . [Accessed 9 June 2015].

[91] J.Nieminen, T.Savolainen, M.Isomaki, B.Patil, Z.Shelby and C.Gomez, “IPv6 over BLUETOOTH(R) Low Energy,” IETF (Internet Draft), 22 May 2015. [Online]. Available: https://tools.ietf.org/html/draft-ietf-6lo-btle-13. [Accessed 02 June 2015].

[92] ZigBee Alliance, “ZigBee IP and 920IP | The ZigBee Alliance,” [Online]. Available: http://www.zigbee.org/zigbee-for-developers/network-specifications/zigbeeip/. [Accessed 02 June 2015].

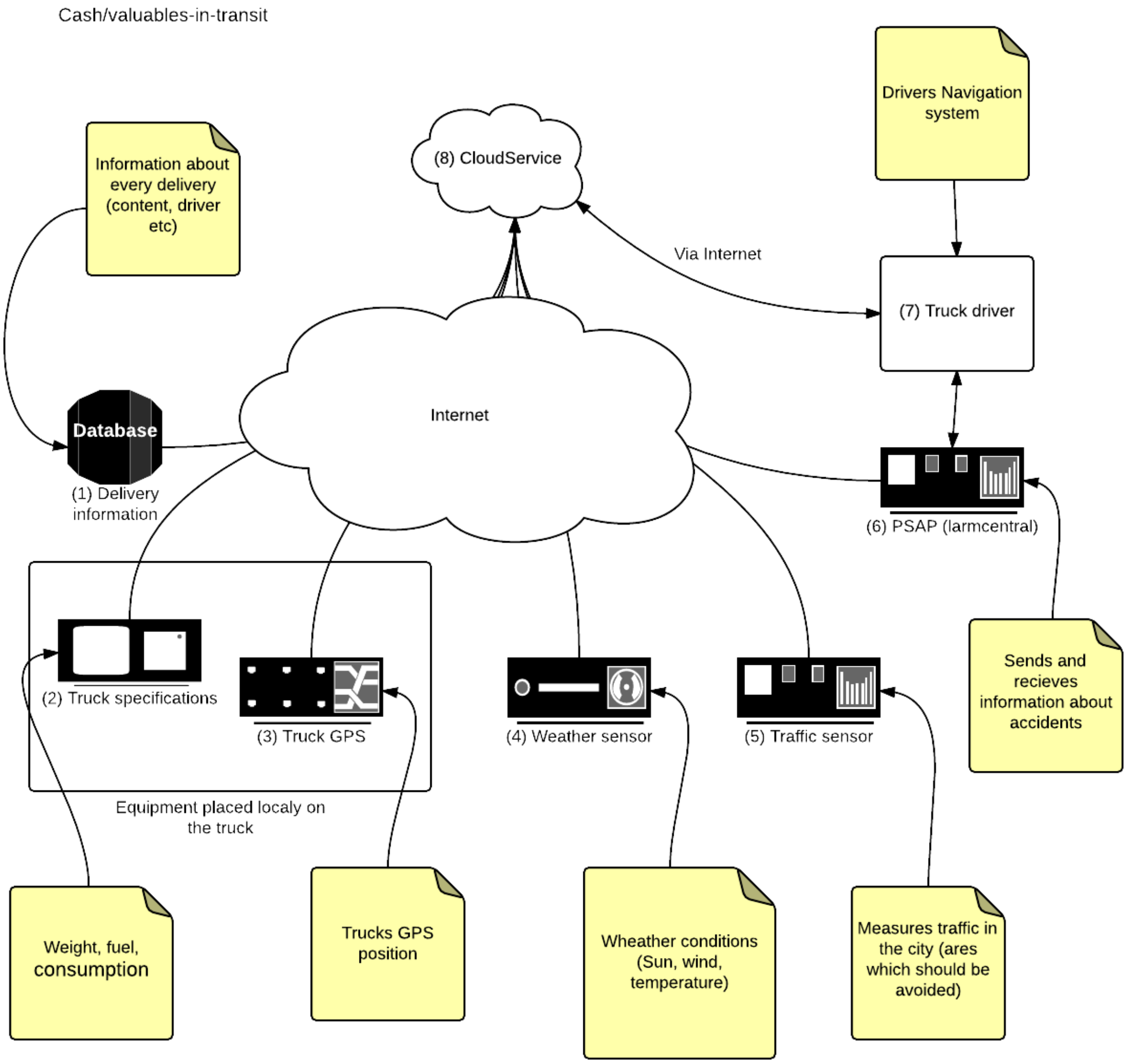
[93] R.Singh, “IKEv2 based lightweight secure data communication,” IETF (Internet Draft), 21 March 2014. [Online]. Available: https://tools.ietf.org/html/draft-amjads-ipsecme-ikev2-data-channel-01. [Accessed 9 June 2015].

[94] "Creative Commons — Attribution 3.0 Unported — CC BY 3.0," Creative Commons, [Online]. Available: http://creativecommons.org/licenses/by/3.0/deed.en_US. [Accessed 13 April 2015].

[95] S. Harris, All In One CISSP Exam Guide, Fifth Edition, McGraw-Hill Companies, 2010.

[96] L. Astrand and T. Yu, “Deprecate DES, RC4-HMAC-EXP, and Other Weak Cryptographic Algorithms in Kerberos,” RFC 6649 (Draft Standard), July 2012. [Online]. Available: https://tools.ietf.org/html/rfc6649. [Accessed 9 June 2015].

Appendix A: Presented IoT Environment


Cash/valuables-in-transit
Information about every delivery (content, driver etc)
(8) CloudService
Drivers Navigation system
Via Internet
(7) Truck driver
Internet
Database
(1) Delivery information
(6) PSAP (larmcentral)
(2) Truck specifications
(3) Truck GPS
(4) Weather sensor
(5) Traffic sensor
Sends and recieves information about accidents
Equipment placed localy on the truck
Weight, fuel, consumption
Trucks GPS position
Wheather conditions (Sun, wind, temperature)
Measures traffic in the city (ares which should be avoided)

Appendix B: Description of IoT environment (in Swedish)

Realtidssystem för värdetransporter. Systemet (molntjänsten[8]) ska förse föraren[7] av en värdetransport med den mest optimala rutt vad gäller effektivitet och säkerhet. Transporten ska med hjälp av information från vädercentral, trafikcentral, larmcentral samt transportens egna specifikationer i realtid avgöra vilken väg som är bäst i avseende till trafikförhållanden, väderförhållanden, olyckor samt vägspecifikationer (max-höjd/ tyngd på vägarna).

Det optimala är om alla del-system är tillgängliga under hela tiden för att kunna göra den perfekta färdanalysen.

Krav (Dessa var menade för Intervju 2):

- Lastbilsspecifikationerna[2] och körningsspecifikationerna[1] behöver vara tillgängliga för att beräkna max körsträcka samt vilka möjliga vägar värdetransporten kan ta.
    - Körningsspecifikationerna[1] behöver bara skickas en gång till molnet. Då informationen är känslig (destination och innehåll) så måste den vara konfidentiell- och integritetsskyddad.
    - Lastbilsspecifikationerna måste även de vara konfidentiella och integritetsskyddade.
    - Lastbilens GPS måste vara konfidentiell och integritetsskyddade.
        - För att inte människor ska kunna spåra värdetransporten.
- Väder[4] och trafikdata[5] och koppling till larmcentral[6] måste vara integritets skyddat för att kunna förlita sig på informationen. Dessa sköts av tredje part och enheterna behöver därmed inte säkras, endast kommunikationen.
- Kommunikation med värdetransportföraren måste alltid vara tillgänglig (viktigaste delen av systemet) så att ändringar under färd alltid kan göras.

Tekniska krav (Intervju 3):

Förtroendemodell:

- Autentisera alla enheter i systemet för att kunna lita på vem som skickar och tar emot data. Hur kan detta göras? Certifikat?

Databas[1]:

- Data ska lagras och skickas säkert (konfidentiell och integritets-skyddat). Leveransinformation skickas endast en gång per leverans till molnet. I molnet behöver det data kunna analyseras i klartext.

Lastbilsspecfikationer[2]:

- Realtidsinformation som ska vara konfidentiellt och integritets-skyddat. Skickas till molnet.

Lastbils-GPS[3]:

- Realtidsinformation som ska vara konfidentiellt och integritets-skyddat. Skickas till molnet.

Väder[4]- och trafikdata[5]:

- Kräver integritetsskyddad kommunikation.

Larmcentral[6]:

- Kräver konfidentiell- och integritetsskyddad kommunikation.

Lastbilschaffför[7]:

- Kräver konfidentiell- och integritetsskyddad kommunikation.

Molntjänsten[8]:

- Data ska lagras och skickas säkert (konfidentiell och integritets-skyddat).
- Centraliserad loggning för spårbarhet i systemet.

TRITA-ICT-EX-2015:93